ASPECTS OF RISK MANAGEMENT IN DEREGULATED ELECTRICITY MARKETS: STORAGE, MARKET POWER AND LONG-TERM CONTRACTS

A THESIS SUBMITTED
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN MANAGEMENT SCIENCE
IN THE UNIVERSITY OF CANTERBURY

by

Stephen R J Batstone

University of Canterbury

2003

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support of many people.

I would like to thank God, for His guidance and strength.

The support, patience and encouragement of my supervisor, Dr Grant Read, has been fundamental to the completion of this thesis. Many staff of the Department of Management, University of Canterbury, have also contributed to this work, particularly Deb Chattopadyay, who gave advice during some very difficult phases of the analysis.

Thanks to my parents, who have never wavered in having faith in me.

Last, and by no means least, heartfelt thanks to my wife Kat, who probably suffered the most for this!

ABSTRACT

This thesis considers the interaction of storage, gaming and forward contracts as mechanisms of risk management in a deregulated electricity market. To date, analyses of imperfect electricity markets have established the "tactical" effect of contracts, that under certain conjectural assumptions, forward contracts have a significant impact on the gaming behaviour of dominant firms in the spot market. However, little work has considered contract strategy, incorporating factors such as the feedback of spot market behaviour on contract market equilibria. This thesis assesses the risk that market participants are exposed to, establishes a measure of risk, and analyses the role that forward contracts play in hedging that risk. Finally, assuming participants are averse to financial risk only, a multi-period model will be provided that examines whether incentives exist for dominant suppliers to use their market power to amplify the risk faced by consumers in order to increase profit through contract revenue.

TABLE OF CONTENTS

LIST OF FIGURES

1

INTRODUCTION

1.1 Motivation

The past decade has seen the progressive deregulation of energy industries worldwide. Electricity generation firms, previously regulated and controlled by government, are now exposed to market forces, implying the transition from a cost minimisation to profit maximisation perspective. It is in this broad decision-making context that analytical research can make significant contributions.

Firms with a dominant position now face incentives to profitably control market outcomes. Historically, the electricity industry has been a natural monopoly, due to factors such as the cost of building transmission networks and the need for a single centralised control centre. Deregulation in most countries has seen the operation of the network separated from the newly formed generation companies, removing a large barrier to the entry of new firms. Still, many electricity markets remain characterised by a small number of large generation firms, and even where a larger number of firms exist, market dominance can still be evident locally or at particular times. Thus the market power models of traditional industrial economics become relevant to an analysis of this situation, and often need to be extended to incorporate the special characteristics of electricity markets.

In any market which is exposed to uncertainty, there is potential risk for its participants. Clearly, generation firms using a fuel whose availability is uncertain face risks in times of fuel shortage. This is particularly true of hydroelectric generation, which is reliant on variable and somewhat unpredictable inflows to operate. In addition, the process of deregulation and the transition to profit maximisation, have exposed all market participants to new risks. While firms were previously concerned with, for example, technological efficiency in a relatively predictable environment, now market share, regulatory policies, and input and output prices are relevant to decision making. In many cases, these variables are largely uncertain and may lead to risky positions for the firm, regardless of whether it is a producer or consumer of electricity.

In order to manage this uncertainty, firms may avail themselves of hedging strategies. Firms with uncertain inputs will utilise storage facilities (e.g., fuel stockpiles or hydro reservoirs) in order to smooth out the variability in supply quantity or price. The presence of storage reservoirs allows the generation firm to absorb input variation and shift water from one period to the next. Hence the ability to mitigate the financial risk of uncertain inflows with storage capabilities becomes an important issue.

Additionally, financial instruments have been developed for electricity markets, of similar forms to those developed in traditional commodity markets. These instruments have enabled market participants to hedge financial risk, but have increased the complexity of decision-making by firms who desire, or are required[1], to simultaneously operate in a market for financial contracts, as well as a physical market for electricity.

1.2 Example: New Zealand Electricity Industry

Since the late 1980s, the New Zealand electricity industry has undergone a process of deregulation. In fact, NZ is often cited as one of the more "successful" deregulation experiments. In 1999, the state-owned enterprise (SOE) Electricity Corporation of New Zealand (ECNZ) was split up into three smaller generation firms, with the major goal of fostering competition and entry by reducing the market share, and thus the dominance, of each SOE in the market. The three stated owned generation firms each had a share of NZ's electric capabilities, and had total market shares of between 15% and 30%, still enough to exert some degree of influence over each other, and the market outcomes. The remaining generation assets in NZ are owned by one large privately owned firm (25%) and a collection of smaller companies.

Approximately 65% of New Zealand's electricity demand is generated from hydro sources, the remainder from thermal generation. The hydro generators' storage capacity is sufficient for only 12 weeks of average inflows, or 6 weeks of average electricity load, making the market relatively sensitive to both seasonal and unpredictable hydrological variations. New Zealand is a temperate climate, with a significant proportion of precipitation falling as snow during the winter months, meaning that valuable potential inflows are locked up in the snowpack, often when it is most needed.

Two dry years in the last decade have caused reservoirs to fall to critically low levels. Hydro firms have responded by significantly reducing generation. Although this may just be the result of the firms' prudent management of their inflow risk, the withdrawals, combined with inelastic electricity demand, and New Zealand's reliance on hydro production, had a dramatic impact on the market. For example, in the winter months of 2001, significant reductions in hydropower drove electricity prices up to four times their normal level for that time of year. Hence, this behaviour has also drawn criticism that it is actually reflecting the generators' use of their dominant market position.

The volatility of the spot price, driven by these hydro variations, would appear to make long-term financial contracts between sellers and purchasers of electricity attractive. In New Zealand, the forward "market" is not yet developed to the extent that standardised contract forms are traded in large quantities on a futures exchange (although a futures market exists, but is not used to any great extent by market participants). Hence the

[1] In some countries (for example, the United Kingdom) many firms have been required through the deregulation

majority of contractual arrangements are settled by a process of direct negotiation between buyer and seller. The respective risk preferences of each party, their expectations of electricity price behaviour, and their reliance on the cost of, or revenue from, electricity will determine if an agreement can be reached. The electricity crisis of 2001 crippled one major retailer who chose not to accept a contract offer by a large hydro firm, prior to the crisis becoming evident.

1.3 Outline of Thesis

This thesis will address aspects of the risk management decision problem faced by the generator, particularly

- how to manage uncertainty on the input side,
- how to manage interaction with their competitors, and
- how to manage the dynamics of contract negotiation.

None of these problems are new - a vast body of literature exists in the areas of reservoir management, game theory and financial risk management. However, it is naïve to consider each in isolation. While these aspects are, individually, relatively well understood in the literature, the synthesis of them is not.

This thesis will develop a high-level framework that brings cohesion to the three aspects listed above, in the context of the decision problem faced by the generator. Furthermore, an analytical model will be developed in order to test the hypothesis that long-run market equilibria exist for this problem. While numerical results will be presented that address this hypothesis, we do not intend to model a particular market, or provide empirical evidence of particular strategies being employed by any generator.

process to sell a certain quantity of these financial contracts, known as "vesting" contracts.

The first section of this thesis develops a framework for the decision problem. This includes explanations of the three individual aspects of the problem, and the interactions between them in the context of a review of the relevant literature.

The second section develops a mathematical model of the decision problem. This considers two broad aspects. First, we will empirically assess the risk faced by generators as a result of inflow uncertainty, and the risk faced by consumers of electricity as a result of load and spot price uncertainty. The latter will allow us to define utility-maximising load, and contract demand, for these consumers. Second, we will construct a hypothetical concept of equilibrium, which will allow us to derive conditions for the existence of physical (spot market) and financial (contract market) equilibria, for a profit-maximising dominant generator facing an uncertain input to generation. We will look at both the situation where the generation firms manage their contract and spot positions separately, and when they anticipate the impact of their spot behaviour on contract prices. While providing insight into the incentives acting on the firm in each scenario, the conditions are difficult to solve analytically. Hence the model is solved numerically as a non-linear program, and results obtained for a reasonable range of important parameter values.

Section 1 of this thesis will proceed as follows:

- Chapter 1 is this introduction.
- Chapter 2 discusses the aspects of risk faced by electricity market participants that are pertinent to this thesis.
- Chapter 3 proposes a framework for exploring the issues faced by a supply-side decision maker in managing these risks.
- Chapters 4, 5 and 6 provide a literature review and discussion of market power, contracts, and commodity storage.
- Chapter 7 presents a review of analyses that have synthesised two or more of these. It concludes by proposing research questions for the thesis.

Based on the modelling assumptions proposed in Chapter 7, the outline of Section 2 is:

- Chapter 8 uses an existing model of an imperfectly competitive mixed hydro-thermal market to examine the impact of profit maximising reservoir management and gaming on the degree of profit risk the generation firms are exposed to.
- Chapter 9 provides an analytical model of risk management behaviour by purchasers of electricity, and from this develops a demand curve for contracts.
- Chapter 10 introduces the concept of a long-run equilibrium, provides a multi-period model of the spot and contract market, and describes equilibrium conditions under the assumption that generators behave in a short-run profit maximising manner. This model includes uncertainty on the supply side, in the form of input cost variations.
- Chapter 11 extends the model by allowing generators to anticipate the effect their spot behaviour has on consumers' perceptions of risk, and thus their hedging behaviour. We investigate whether incentives exist for generators to magnify input cost variations in their spot market behaviour, in order to manipulate contract market outcomes to their advantage.
- Chapter 12 presents numerical solutions to the models, in particular, to establish whether firms operating under the "destabilisation" incentives can find stable spot and contract market equilibria.
- Chapter 13 draws conclusions and proposes areas for future research.

2

RISK IN ELECTRICITY MARKETS

2.1 Introduction

Chapter 1 outlined the context in which the thesis will consider the decision problem faced by a supply-side player. To recap, the significant characteristics of interest here are:

- A small number of supply firms, with more than one possessing a degree of market power.
- Generation is a mix of hydro and thermal, with hydro generation a significant contributor to total system load.
- The hydro firm has a moderate capacity to store variable inflows.
- The opportunity exists for long term contracts to be sold by supply firms to retailers or end-users of electricity

The analysis of risk is important to decision making under uncertainty. This necessitates an understanding of the various methods of defining and measuring risk, and of the various models of how decision makers respond to risk when making choices in the context of uncertainty. This thesis is not intended to be a critique of decision models involving risk, however, a summary of the major models can be found in Appendix A, and will be drawn on as required.

Appendix A also highlights the variety of ways in which risk is defined and measured. Perhaps the broadest of these is provided by the Oxford Dictionary, which defines risk as "the chance or possibility of loss or bad consequence". In any deregulated environment, firms faces a wide range of "risks", under this definition, at various levels of operation and over various time frames (Figure 2.1). Section 2.3 will present a discussion of the more general aspects of risk relevant to an electricity market participant, but ignored in this study. Many of these "business" risks are common to most companies, and an analysis which attempted to include all of them would be too complicated. It is also difficult to reconcile many of them with our chosen measurement of risk - the statistical variance of a distribution of outcomes. While we acknowledge that these risks are important, they are left outside the scope of this thesis. However, brief reference will be given to the literature that deals with them, especially in the context of a deregulated electricity market. This section includes a discussion of risks faced by two other important market participants, potential entrants and transmission and distribution companies.

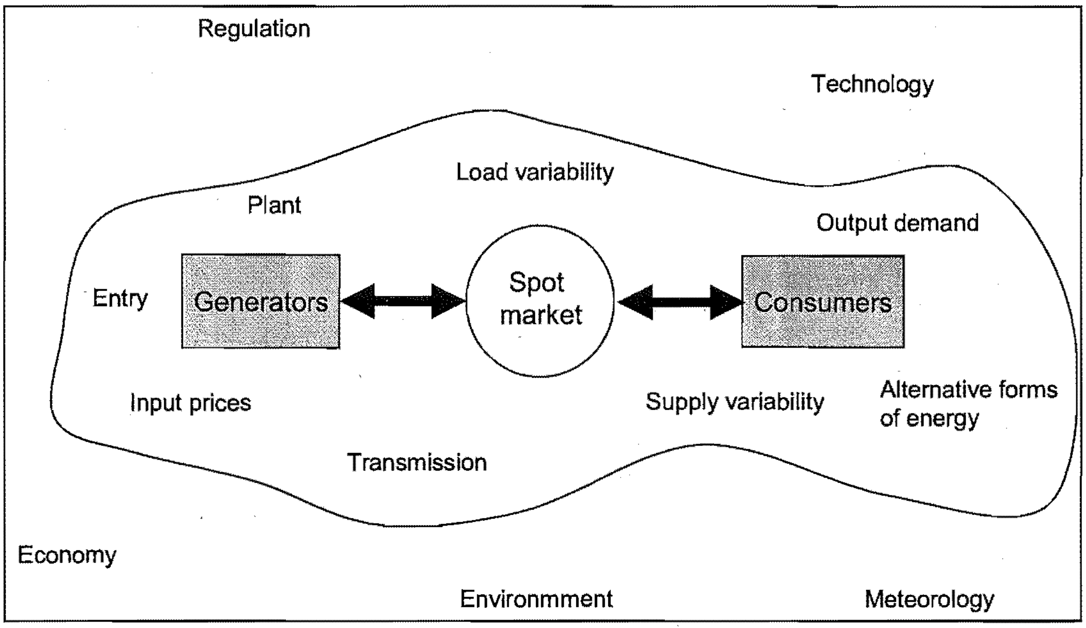


Figure 2.1 Electricity market risks

The chapter will then identify the risks that are relevant to this thesis, and how they will be measured. Section 2.3 describes our choice of risk measurement, i.e., the variance of profit, and then presents a discussion of important driving factors of this risk, namely price and quantity risk. The discussion is presented from the perspective of three types of market participant: supply firms, large industrial users of electricity, and electricity retail companies. This section identifies which particular aspects of price and quantity risk will be used for the remainder of this thesis. While the analysis presented in this thesis is taken from the perspective of a supply side firm as described above, risk measures must also be identified for the demand side firms who will desire to hedge using long term contracts.

Section 2.4 summarises the discussion of risk and motivates a structured approach to risk management for these market players.

2.2 Business Risk

There exist a range of risks that are not particular to an electricity company, but are faced by any firm operating in a modern economy, such as:

- Economy risk: which could include economic growth, interest rate risk and exchange rate risk on foreign investments.
- Credit risk: The risk that debtors will default on their loans.
- Legal compliance: Broader than the regulatory and environmental risks discussed below, this includes issues such as compliance with workplace safety regulations, employment-related risks and other activities of the company that are affected by legislation.

Some of these aspects may in fact be driving factors of the risks discussed in Section 2.3 (for example, economic growth and interest rates may drive demand variation in the long term). However, due to their complex nature (which is really the realm of macroeconomics), such relationships will be ignored.

We now focus on a range of risks that hold particular importance to participants in an electricity market.

Regulatory Risk

Larsen and Bunn (1999) suggest a significant issue for firms in the new market is regulatory risk, which is relevant on two major fronts:

- **Transitional regulatory risks**: Many controls on prices and investment behaviour, for example, are intended to aid the market in making the journey from fully regulated to fully competitive. Such controls are quite common when previously government owned utilities are forced to divest some or all of their generation assets. The risk associated with these controls relates both to the market behaviour they allow or prevent (for example, price caps restricting a firm's ability to pass on high costs), and the uncertainty surrounding the regulator adding and/or removing controls, and how the market adapts to the new environment.
- **Antitrust risks**: Firms must be not only concerned with how their competitors will react to their strategies, but also whether the market regulator will view any

action as potentially anti-competitive. The major risk to firms is to invoke disciplinary action from the regulator.

Regulation may also occur in response to one-off situations that endanger the stability of the market, for example shortages arising from unexpected high demand, long-term outage of transmission or generation plant, or the exit of significant market participants. Governments believe it is their role to ensure the availability of electricity as a public good, and may intervene in these situations. Regulatory bodies may also provide incentives for investment in generation capacity to increase certainty in the markets ability to meet demand in the long-term, if they perceive the market signals for such expansion are not adequately provided.

Often such issues can only be effectively investigated with simulation models. It is extremely difficult to model such risks using analytical methods. While most regulatory actions can be represented quantitatively in a model, the functions representing the behaviour of the regulator in response to market scenarios or firm behaviour may not be "well behaved". Determining the likely outcome of these scenarios can only be modelled by a series of "what-if?" analyses.

Environmental Risk

One increasingly important area of regulation is with regard to the environment. Fossil-fuel fired electricity plant is a significant contributor to greenhouse gases, and these generation methods have come under increasing scrutiny from governments worldwide, especially those countries in Annex-1 as defined by the Kyoto Protocol[2]. For some countries, this agreement requires significant and costly (to the country) reductions in output from fossil fuel fired generation plant.

Given the significance of electricity generation to carbon dioxide emissions, authors have begun to analyse future scenarios for the restriction and retirement of fossil-fuel generation plant. Rothwell (2000) suggests that the required reductions in carbon dioxide

[2] The Kyoto Protocol was the world's first international government agreement on reducing carbon dioxide emissions, signed initially in 1997. Essentially, the Kyoto deal requires most OECD countries to reduce their carbon emissions to close to 1990-levels.

emissions may be so dramatic that relatively expensive nuclear plant may be rescued from being mothballed. However, since most governments are yet to carry out the full extent of their obligations under Kyoto, and the agreement itself is still shrouded with controversy, these analyses have done little to remove the uncertainty of future scenarios. For a comprehensive investigation of the likely implications of Kyoto, see Weyant and Hill (1998).

The ambiguity surrounding Kyoto is set against a wider backdrop of environmental regulation. Many commentators pointed to the strict environment restrictions on new power plant as being a major factor in the California power crisis of 2000/2001. Andrews and Govil (1995) highlight the risk associated with environmental regulation, arguing that environmental regulations are a reactively driven "moving target", and result in suboptimal compliance and investment strategy.

Entry/Exit Risk

While assessments of risk in the short term can be based on knowledge of the existing market structure, over a longer time scale, average prices in the industry may lead to a change in the market structure, or, more specifically, the entry and exit of firms.

From the incumbents' perspective, the possibility of entry introduces an extra dimension to price and quantity risk. Entry may force prices down as the entrant competes for market share. Much of this depends on the marginal cost of the entrant, and where in the merit order they are likely to generate. Expensive thermal or peaking plant may be called on less often due to the additional capacity in the system. Baseload incumbents may not have their output reduced by entry, but profits in high demand (and therefore high priced) periods may be reduced if cheaper peaking plant becomes available to the system.

The entrant's owners must be convinced that there is enough "space" in the system to ensure long-run viability. Kriebel and Hornstein (1999) points out that the financial success of electricity plant is dependent on "its ability to produce electricity at a total cost to the generator, including all debt service, which is, on average, less than that which the market will pay for such electricity". Thus ensuring that **average** prices will recover total costs may be sufficient for cheaper plant that will be operating all the time, and such entrants may prefer a stable market environment.

Peaking entrants, on the other hand, are not so much concerned with the average price (which may be below their marginal cost) and could in fact prefer highly volatile prices, as this increases the chances that they will be called upon to generate.

In either case, the risk associated with entry will depend on the incumbents' reaction to entry (for example price wars) and the nature of the system (a hydro dominated system with significant seasonal effects introduces greater opportunities for expensive generation in the low-inflow periods).

Here the concept of a price duration curve (PDC) becomes important. A PDC shows the number or proportion of hours that the price exceeds a certain level. For a high priced thermal, the question of exactly when they will be called on to produce is not so important as whether the number of hours they spend generating at a price greater than or equal to their marginal cost is large enough to ensure the viability of operation.

For cheaper "baseload" plant that will be dispatched each period, the assessment of risk reduces to the uncertainty surrounding the market price that will prevail, and whether the price-marginal cost differential is large enough to recover the fixed costs over a reasonable period of time. More expensive plant face the additional uncertainty of not knowing when demand conditions, or competitors' offers will be such that their operation will be accepted and profitable.

In addition to the risks associated with the operation of the new plant, the process of building new plant for incumbents or entrants has uncertainties that have significant financial consequences. Kriebel and Hornstein (1999) provides a framework for the analysis of construction and technology risk associated with new plant (project risk).

Transmission Risk

A fifth player could be considered in this analysis. In most deregulation cases, the transmission and distribution (T&D) role has been assumed by an independent firm, often government owned, and remains a natural monopoly. This firm must manage the risks associated with system reliability and capacity investment to meet growing demand, in the context of uncertainty. While this area of electricity risk will not be considered further here, the reader is referred to Hoff (1997) for a distributed resources approach to T&D planning and operation under risk. Burchett and Tummala (1999) details a "structured approach to enumerating and understanding potential risk factors and assessing consequences and uncertainties" for a transmission line construction project.

Other Generic Risk

Given the common theoretical definitions of risk proposed in Appendix A, the predictability, and thus the available history, of uncertain variables plays a vital part in gauging the magnitude of risk a firm might be exposed to in the future. This could be regarded as one of the most significant risks that all players in a deregulated market face. Most electricity deregulation experiments are relatively recent, making even the process of defining distributions for the evaluation of risk a difficult process. Strategic behaviour, demand elasticities and the effect of long term contracts have to be understood as well as modelled, and a lack of available data only compounds this process. Larsen and Bunn (1999) motivates the use of simulation as a decision making tool, given the lack of available data.

The authors also provide some additional industry-wide risks prevalent in the new market structure, such as the corporate risk associated with the knowledge base of the firm's executives as it makes the transition from technical and administration-based management to market-focused objectives. Also relevant to this "corporate risk" are the implications for the sale, purchase, merger and acquisitions of electricity firms in the capital markets:

> *"Not only can a company no longer take prices and its own customer base for granted, it is also subject to the forces of the capital markets and cannot even be sure who its owners may be from year to year."*
> Larsen and Bunn (1999) p340

2.3 Market-related Risk

We now turn our attention to those risks associated with the firm's operation in the spot and contract markets, that in turn drive profit. The ability to manage **profit risk** is the primary focus of this thesis.

Profit, at its simplest level, is driven by four factors:

- The price and quantity of inputs used in the production of the good, determining total cost

- The price and quantity of the good sold in the market place, determining total revenue

Each of these four factors will be dealt with for both supply and demand side firms. Under the modelling assumptions outlined in Chapter 1, some of these will become irrelevant, as discussed below. Firstly, though, we briefly defend our choice of risk measurement.

2.3.1 Measuring Risk

A cursory glance at the survey of literature presented in Appendix A reveals that a number of measures of risk have emerged over the years. To briefly summarise, risk arises from the uncertainty of future events. Traditionally, this has led to risk being measured with a variable relating to the distribution of possible outcomes. This is complicated when risks are intangible, not easily represented quantitatively, or not distributed continuously, such as when the distribution includes one-off events that aren't part of the "normal" range of outcomes observed, such as sharemarket crashes, strikes and other catastrophic occurrences. However, since this thesis will concentrate on those aspects driving profit risk outlined above, i.e., prices and quantities, outcomes naturally take numeric values, and in most cases result in "well behaved" distributions.

Since we are concerned with profit risk, we require a measure of the risk in a given distribution of profit. A popular measure of financial risk is the statistical variance of the profit distribution. A distribution with a high variance is, by nature, more variable, and therefore the outcome value was less certain. Traditionally, it is assumed that rational, risk averse decision makers prefer less variability to more.

While some authors have suggested other moments should be used (skewness and kurtosis) to obtain a more complete picture of the distribution, variance has been retained mainly because of its computational convenience, and its wide use, and subsequent intuitive appeal to managers in general.

Variance has faced its most significant criticism from the proponents of downside risk. Downside risk was argued to be a more accurate depiction of the widely held belief that risk was really only those outcomes that affected the decision maker adversely. The Oxford Dictionary definition of risk, quoted earlier in the chapter, supports this claim. Hence the notion of downside risk was developed to only measure that part of the distribution that is below a target or expected value. Variation above such a target was not believed to be risk, but rather an unexpected pleasant surprise.

While we will initially refer to the statistical variance as our measure of risk, the extension to a general form of downside risk that retains the analytical convenience of variance is not complicated, and, in fact, will be used later in this thesis.

A caveat must be noted here. It is tempting to imply from the above discussion that all variability is risk. This is clearly not the case. If, for example, prices follow an entirely deterministic pattern over time, a firm will know in advance what the outcome in a given period will be. Even if some of these outcomes are "bad" for the firm, we would expect that a rational decision maker would adjust their firms operations in advance to account for such outcomes. Whether this process of adjustment is hedging, and hence the bad outcome is still risk, is an area of philosophical debate that will not be addressed here.

We will assume that risk only arises from those outcomes that are unexpected. If part of the variability of any driving factor is predictable, then only the residual "volatility", where the outcome of the variable is different from that predicted, will be considered risk.

Hence our ability to measure risk depends on how well we can separate out this volatility from the overall variability of any given factor. If we are to ignore predictable patterns of variability, and measure the variance over all individual observations, then we would potentially overestimate the true risk inherent in a firm's position in the market. On the other hand, we could rely on long-run averages to describe risk. For example, if the annual average spot price is greater than expected, then price observations during the year must have departed from their predicted path. This, however, has the danger of underestimating risk, since an average muffles the effect of an individual outcome, which, even by itself, might have dire consequences for the firm. An extremely high spot price may only occur in one week of the year, but if it comes at a time when a net purchaser of electricity is highly exposed (i.e., holds a low level of contracts relative to electricity requirements), this situation could be catastrophic for the firm.

While we will proceed with using the statistical variance of a distribution as a measure of risk, how it is used will depend on the situation, and will be addressed as it arises. In general discussions involving risk, "volatility" or "risk" will be used to represent that part of the natural variability which is not predictable, rather than the overall variability of the factor.

2.3.2 Demand Risk

Two important classes of consumer exist on the demand side of electricity markets: those who purchase electricity directly off the spot market, and those who purchase through a electricity retailer[3]. Those who purchase in the former manner tend to either be large industrial consumers or the retail companies themselves, while the latter is made up of households and small to medium sized businesses. In New Zealand, 1998 figures showed that approximately 70% of total electricity consumption was distributed through retail companies, while 22% was purchased directly off the spot market or through hedge contracts by less than 10 large industrial companies[4]. Co-generation methods represents the remaining 8% (Commerce (1999)).

---

[3] A retailer could, in fact, be a generator, if the particular market allows vertical integration. This issue will be dealt with in Chapter 9.

[4] The remainder involves smaller industrial firms that owned on-site co-generation plant.

These two groups of large spot-purchasers (retail companies and large industrial firms) face similar input price and quantity uncertainty (i.e., the cost and availability of electricity), but different output price and quantity uncertainty.

Input Risk

Given that transmission and plant availability issues are ignored here, we can assume that the quantity of electricity these firms require will always be available from the spot market. Clearly, both types of consumer face input price risk, as we assume neither retailers nor industrial firms have any influence over the spot price. If risk averse, both will find risk management strategies that reduce their exposure to the electricity spot price attractive.

These firms possibly differ with respect to the magnitude of risk they are exposed to as a result of electricity spot prices. As opposed to retailers, where electricity is the product they sell, electricity is only an input to a complicated production process for industrial firms, and the degree of profit risk that can be attributed to input price volatility will largely be determined by the proportion of profits attributed to total electricity cost. Obviously, if electricity is a relatively insignificant cost, price variations will have little overall effect on profits. However, in many cases, the fact that the firm purchases its electricity directly from the spot market, rather than through a retailer, would imply that electricity cost is a significant component of profits.

Output Risk

Industrial consumers

Not only do industrial firms face uncertainty in input prices, they also face, to differing degrees, volatility in the quantity and price of their output. Periods of high product demand, or low product prices, coinciding with high electricity prices increase profit risk, and vice versa. Many firms (for example, a large snowmaking skifield) face seasonal demand which results in significant electricity requirements during a time of year when prices in a hydro dominated electricity industry are at their highest. Unpredicted movements in quantity and price at these times are often more critical than other periods in the year.

In summary, the following factors will have a significant effect on these firms' output risk:

- The volatility of demand for their product
- The ability to control output quantity and price. By virtue of their size, it is likely that such firms will be significant players in their respective product markets, even if they hold no such market power in the electricity market
- The ability to shift product demand between periods – i.e., positive and negative (back-ordered demand) storage

Electricity Retailers

As an on-seller of electricity, retail companies are not concerned with physical assets, environmental issues and other capital-intensive factors faced by generation firms. Their revenue is derived from the sale of electricity to a customer base, which through retail deregulation is no longer stable over time. Costs are incurred directly from the purchase of an uncertain quantity of electricity off the spot market, or through hedging arrangements with suppliers. Hence the major source of profit risk to a retailer is the combination of:

- Load Risk: Uncertainty about the quantity of electricity demanded by a given customer base in real time, and changes in the size of the customer base in the mid to long term, as customers respond to price and non-price competition between retailers, or even alternative forms of energy, e.g., gas and solar power.
- Margin Risk: The retailer's margin is the difference between what it paid for electricity and what it receives from its consumers. Margin risk could be defined as the retailer's ability to, in the long run, on-sell electricity to end-users for more than it paid for it.

As a consequence of filling the role of on-seller to large customer bases, retail firms have become "mega-purchasers" of power, and thus more exposed to price risk than any other part of the industry Gersten (1999).

Short-term **load risk** arises from the fact that many individual consumers are making decisions on load usage in response to a large number of factors. While the vast majority of consumers purchasing from retailers don't observe the spot price for electricity (see below), day-night differentials in fixed-price electricity sale agreements do lead to a degree of load-response to the higher- and lower-priced periods, thus reducing the risk retailers are exposed to.

While electricity end-user profiles help predict variations across a day and week, uncertainty still arises from other exogenous factors, such as weather. Longer term changes in average load are much more difficult to predict, as they are often driven by unpredictable factors, such as the economy and technology (ignored here)[5].

Margin risk is inextricable from the input price risk faced by the retailer. Retail electricity firms may have a proportion of their end-user demand on fixed price contract, at least in the short-term when the retailers have no ability to alter electricity rates. Not only does this make the process of setting such rates very difficult, but it is also compounded by the fact that the load itself will be inelastic with respect to the spot price. While retailers may sign fixed input-price contracts with generation firms, margin risk may still exist. If this hedge is insufficient to meet the retailer's consumer load, particularly likely in periods of high demand, the shortfall will have to be purchased off the spot market, at uncertain prices. These extra purchases may come at a time when the spot price is high, in response to the high levels of demand. Conversely, if the amount of the hedge is surplus to requirements, the excess must be sold back to the spot market, possibly at a loss on the contract price paid for it.

While the obvious risk associated with the ease of "switching" (the process by which a consumer changes electricity retailer) is for the retailer to lose customers, the upside also constitutes a risk. Retailers compete for customers on a range of price and non-price competitive strategies. If the success of a marketing initiative is greater than expected, a retailer can be exposed to more price and load risk than it was initially anticipating[6] (Keers (2000)).

2.3.3 Supply Risk

Input Risk

Firms who operate thermal plant fired by coal, gas or oil must purchase those fuels, either at an uncertain spot price, or through hedge contracts. While the physical marginal cost of hydro plant, and thus price risk, is negligible, the ability to operate is determined by hydrological inflows, representing a significant quantity risk. In many climates weekly inflows are highly uncertain and seasonal, and often the periods with the greatest demand coincide with the lowest inflows. In temperate latitudes with alpine regions, rainfall is trapped in the form of snow for many months of the year, and arrives into reservoirs in a very short space of time as the snow melts in spring. Firms use storage reservoirs to "shift" water between periods to minimise this aspect of quantity risk, in the same way as purchasers of fuels for fossil-fired plant store their input, in order to hedge fuel price uncertainty.

Output Risk

In terms of the generation firm's output, quantity risk comes in two major forms:

- Inability to generate or transmit a desired quantity of electricity to the market.
- Uncertainty of (residual) demand for electricity.

Transmission issues and plant outages are pertinent risks for an electricity generator in terms of an inability to capitalise on favourable market conditions. Additionally, being in

---

[5] One only needs to examine recent power crises to appreciate the implications of uncertain load levels. The California crisis of 2000/2001 was in part caused by unexpected escalating load, as air-conditioning demand rose during an abnormally hot summer. While profiling may have accurately predicted the half-hourly variations in load, the underlying level of load was beyond forecasting. Other medium-long term factors contributed as well, such as the growth over a number of years of electricity intensive computer firms in the area, all insulated from the real spot price of electricity.

[6] Recent events in New Zealand have seen retail firms "close their books" to additional consumers, in an attempt to limit the risk they face in times of high electricity prices.

a position where full plant capacity is not available, or the desired quantity of electricity cannot be transmitted to the market, introduces risks beyond just the loss of potential profits in favourable market conditions. Firms with high levels of hedge contracts may find themselves "caught short" by their inability to generate. If the contracted amount is greater than their own available capacity, then the shortfall must be purchased off the spot market by the generation firm themselves, thus becoming a "net buyer" of electricity[7]. If the spot price is higher than the hedge price, a loss is incurred by the firm. In this way, a hedge may provide certainty during "normal" operation, but can significantly increase losses during periods of outage.

The other significant component of output risk relates to the quantity that will be accepted by the market clearing mechanism. It is more appropriate to consider quantity and price risk together here, as it reflects the risk more accurately.

The total quantity of electricity demanded at any point in time is driven by a large number of independent decision makers, making accurate prediction of loads difficult. However, depending on the relationship between price and quantity (the shape of the demand curve), many of these demand variations may be absorbed by offsetting price variations, and thus having a much reduced effect on profit risk.

However, the variation in demand only represents part of the quantity risk faced by suppliers.

The variation in price and quantity sold into the market is driven not only by demand uncertainty but also uncertainty surrounding the actions of a firm's competitors. In markets where offer curves are submitted to a market clearance mechanism, certainty of dispatch, and which plant sets the market clearing price, become central to an analysis of risk. The risk of submitting an offer at a high price is that the plant won't be dispatched. This in turn is highly dependent on other firms' actions.

[7] These concepts relating to two-way contracts are explained more fully in Chapter 5

A profit maximising firm won't normally[8] offer electricity any lower than marginal cost, so the nature of the generation plant becomes critical. Some technologies (e.g., oil-fired thermal) have significantly higher marginal costs than others (e.g. hydro, nuclear), and have less ability to offer electricity at a price which will almost guarantee dispatch than their cheaper counterparts. This is complicated by strategic effects, where firms with market power choose to offer at prices greater than marginal cost, safe in the knowledge that their significant market share, and the relative marginal cost of their plant to their competitors' necessitates their dispatch in order to meet demand.

2.4 Conclusions

Table 2.1 summarises the aspects of price and quantity risk outlined in the previous section. As discussed there, not all of these aspects are relevant to this study.

ASPECT OF RISK	SUPPLY SIDE	DEMAND SIDE
Input Price	**A**: Cost of inputs e.g., gas, oil	**E**: Spot/Contract cost of electricity
Input Quantity	**B**: Quantity available e.g., inflows	~~**F**: Supply availability~~
Output Price	~~**C**: Demand and competitor response~~	~~**G**: Price of good[9]; Margins[10]~~
Output Quantity	~~**D**: Plant or transmission outage~~	**H**: Demand for good; Demand for electricity

Table 2.1 Risk treated in this thesis

This study ignores the issues surrounding plant and transmission failure, and focuses on the actions of a hydro generator, whose water comes "free", but in uncertain quantities. As shown in Chapter 8, the profit maximising management of storage reservoirs effectively transform the hydro generator's problem into one of uncertain input costs, where the costs are marginal water values. Hence we will generalise the risk faced by

[8] Unless it is over-contracted, and the optimal response is to generate some of the shortfall rather than purchase it off the spot market at the spot price.

[9] For industrial firms purchasing directly from the market

[10] For bulk electricity retailers, who on-sell the electricity to small consumers and households

supply firms to input price risk (i.e., we include box B in box A), to allow thermal firms, facing uncertain fossil fuel prices, to be included in the analysis.

While the form of strategic interaction between the firms is a central theme in this study, we assume that all supply firms know the form of the "game", and thus can predict market outcomes (for given input costs) with complete certainty. The case where a dominant firm is unsure how its competitor will react would be an interesting extension to the models presented here, as would the development of offer curves under uncertainty[11]. Similarly, we assume that firms know the demand curve they face, in both the spot and contract markets.

Hence supply firms in this thesis face risk only in box A, i.e., input price.

By the same reasoning, we assume that consumers and retail firms will always be able to purchase their requirements with certainty, i.e., they are not affected by supply or transmission outage. However, we will model their uncertainty about the spot price (box E).

Since the majority of the analysis in this thesis is performed from the perspective of a generator, detailed modelling of the demand (and inherent uncertainty) faced by retailers and industrial consumers will be aggregated into a (potentially uncertain) load for retailers, and a demand curve for goods produced by industrial firms. So, while retailers face output quantity risk (box H), the margin for electricity, or price for the good, is either ignored, or known with certainty (for a given load). However, they are still exposed to risks in boxes E and H, which are taken up in more detail in Chapter 9.

Having identified the aspects of risk that are relevant to this study, we now proceed to establish a framework to manage this risk, using three important mechanisms available to the electricity generator as modelled in this context.

---

[11] See Anderson and Philpott (2002b)for an excellent treatment of the offer stack construction problem, for a firm that is uncertain about how much of the "stack" will be dispatched by the market.

3

RISK MANAGEMENT IN ELECTRICITY MARKETS

3.1 Introduction

The previous chapter outlined areas of risk for an electricity market participant, and concluded by narrowing the definition of risk for this thesis to financial risk, where risk is represented by statistical variance of a distribution[12].

We now turn our attention to the management of risk. This chapter will introduce three mechanisms available to a electricity market participant to manage risk as defined in this thesis, within a framework which will enable us to examine the roles each of them play, both individually and simultaneously.

This chapter will make brief references to the vast literature on measurement of risk and risk attitudes. A more comprehensive and detailed survey of this literature is provided in Appendix A.

---

[12] As noted there, while this will initially be assumed to represent the entire distribution of profit outcomes, we will later restrict this to measure only those profit outcomes that are "bad" for the decision maker, i.e., downside risk.

3.2 Risk Management and Risk Attitude

Ward (1997) argues that true risk management, by definition is a "systematic and professional approach to improving performance via identification, appraisal and management of relevant threats and opportunities." The "management" Ward alludes to is generally defined as the reduction of the likelihood and/or magnitude of bad outcomes[13].

A significant proportion of the risk management literature has its roots in the insurance industry, and applies to catastrophes, legal and environmental compliance, intellectual property and other infrequent (possibly even one-off) situations which represent significant costs which could be devastating to a firm. The notion of minimising the adverse impact to the firm underlies most risk management approaches relevant to these situations. Techniques used here have been extended to general project management within a firm, especially at the corporate planning level. While this kind of analysis is very relevant to electricity firms, who, for example, face significant and infrequent events such as plant failure, these occurrences fall outside the definition of risk as presented in the previous chapter.

Chapter 2 proposed that the range of profit outcomes, from which risk will be measured in this thesis, can be represented by a probability distribution. It is reasonable to expect that individual decision makers will value each of these outcomes intrinsically, reflecting their attitude to the "riskiness" of the distribution. The process by which they attempt to change the distribution, and/or make the best decision could be loosely termed risk management. The literature presents a variety of ways of representing both the value function and the process of minimising risk (surveyed in Appendix A).

We wish to account for two distinct classes of risk attitude, where risk is defined as variance of the profit distribution, namely:

---

[13] As an aside, Ward goes on to say that many managers view RM techniques as simply avoiding bad outcomes, rather than looking for good opportunities as well.

- risk aversion, corresponding to decision makers who seek risk management strategies which reduce profit variance

- risk neutrality, where decision makers are unconcerned about profit variance

Most traditional categorisations of risk attitude include a third category, namely risk loving, where decision makers pursue strategies that increase the level of variance in their profit. It may seem unusual that a manager of a large firm in an electricity market would fall into the last category. While we later investigate strategies whereby a firm will seek to increase the variability in *other* market participant's profit, we still assume in this study that no firm would wish to increase variability in its own profit. Hence firms will always be either risk neutral or risk averse towards their own profit distribution.

We also assume that these risk attitudes can be incorporated into a single function of risk and return, where the degree of risk aversion can be represented by a single parameter.

3.3 The Risk Management Triangle

We now present a framework to help investigate how these risks will be managed, in the context of an imperfect, hydro-dominated electricity market. It becomes apparent that there are three mechanisms available to the decision maker to hedge the risks involved in such a market (Figure 3.1).

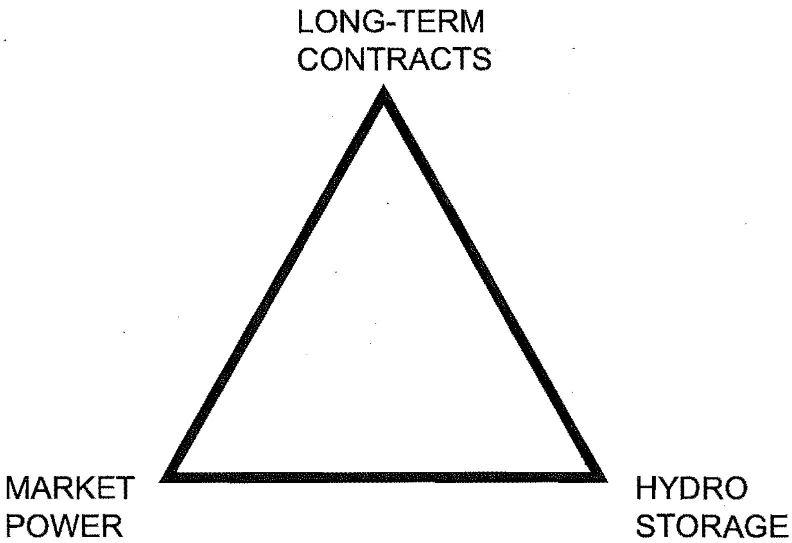


Figure 3.1 Risk Management Triangle

Chapter 2 established that for the purposes of this study, risk is driven largely by the uncertainty surrounding input and output prices and quantities. In particular, supply firms are concerned with variation in input quantities and output prices, while we will model the demand side as being averse to variation in the cost of electricity and in their own requirements (whether load as a retailer or demand and profitability of a commodity as an industrial firm). While this thesis investigates risk management from the perspective of a supply firm, the following discussion shows that the same framework is relevant for demand side firms fitting the same assumptions.

3.3.1 Market Power

Chapter 4 will show that market power is more likely to be evident on the supply side of electricity markets, where large fixed capital costs and increasing returns to scale result in a small number of firms. However, market power issues may be appropriate to the demand side as well. Where a large number of small individual consumers are represented by a few bulk-buying retail companies, who each may be able to influence the actions of other participants, strategic interaction between demand-side participants should be modelled.

While a firm has control over its own behaviour, the degree of market power it possesses will determine whether it treats price risk as within or outside its control. A firm that is a price taker views the variance of these variables as exogenous, and consequently is quite uncertain about profit. Firms who can anticipate the reaction of other participants, through the knowledge of demand and/or supply curves, can reduce price uncertainty and thus risk. A firm that can influence the market outcomes to its own advantage can reduce its own price risk even further (and possibly create it for others).

Market outcomes are the result of all players' decisions. In an oligopoly, an individual firm has no direct control over the pricing and quantity behaviour of other market participants, so uncertainty can still surround final market outcomes. However, firms on the same side of the market may possess an ability to indirectly influence each other through the dynamics of gaming, so the assumptions surrounding competitive conjectures become critical to how much risk is created by firms' simultaneous actions.

A firm that exerts market power can choose to take account of two effects. Firstly, the immediate effect of a particular action in the market is noticed in the spot price for that period. Secondly, the action has a signalling effect, where competitors and consumers interpret the action of the firm as indicative of the firm's broader strategy, and the future of the market. Both these aspects contribute to the present, and future, variability of profit, and thus risk.

3.3.2 Contracts

"Two-way" contracts, or contracts for differences (CfD's) are similar to traditional financial forward or swap arrangements. These contracts allow supply and demand firms to agree on a quantity and/or price of electricity over a given time period, but differ from traditional forward contract arrangements (which will be detailed further in Chapter 5). By locking in fixed prices and quantities, these contracts give an increased level of certainty about overall electricity costs for purchasers and revenues for generators over that portion of total output or total demand.

Further reducing profit risk, supply firms can purchase 'backup' contracts from each other, commonly in the form of financial call options. Transmission or plant outage, high cost of, or lack of, water and other input fuels sometimes restrict a firm's ability to supply their desired generation level to the market, particularly when firms have a high level of supply commitment themselves in the form of two-way contracts with consumers. These contracts allow them to call on other generation firms to supply electricity on their behalf at the agreed strike price.

3.3.3 Storage

Once generated, electricity itself cannot be physically stored. However, the storage effect can be replicated by holding inputs into its production (oil, gas and water) in storage. For hydro, large storage reservoirs give the generation firm the ability to shift water between periods of surplus inflows to times of shortage, in this way acting as a buffer against uncertain inflows. Significant factors here will be seasonal effects, the variance of inflows and the reservoir capacity.

Indirectly, some consumers have the ability to 'store' power also, shifting their demand between periods of high and low electricity costs. Large industrial firms achieve this by either deferring or accelerating production runs.

3.3.4 Interactions

However, these three tools are not mutually exclusive. The risk management triangle includes the interactions between storage, contracts and market power.

Since long term contracts commit parties to certain levels of generation and load, they are generally believed to reduce the firm's incentives to use market power. A large supplier that has sold hedge contracts, for example, has less freedom to profitably vary generation to influence price, than if they had no contracts, as they are committed to produce a certain amount for the contract purchaser. As a result, we would expect to observe average output levels higher when hedge contracts are in place. Additionally, there are less incentives to affect the spot price, since this will (in the short run) not affect the price received for the contract quantity.

Storage may have a similar restrictive role in determining the degree to which market power, or indeed any spot market strategy, is exercised. A chosen level of release can only be executed if the required water is available. Profitable opportunities may be missed due to low levels of water availability.

As market participants observe results in the spot market, they will develop expectations about the future state of the market, described by various parameters such as mean and variance of prices. These expectations become crucial determinants of the participants' willingness or propensity to enter into forward contracts which cover these future periods.

The quantity of contracts sold may influence storage decisions for a hydro generator. The firm may choose to hold stocks of water so that it minimises the risk of not being able to meet its contractual obligations (since it would be required to purchase any shortfall of the spot market at the prevailing price). However, high levels of contracting may lead to low storage levels, as the firm does not have the freedom to hold back generation from the market in times of low inflows.

In the same way, since generation firms with market power choose output levels that are lower than in the competitive alternative, we would expect to observe higher levels of storage among dominant market players. As discussed above, the level of contracting may increase output, and thus result in storage levels somewhere in between the gamed and competitive outcomes.

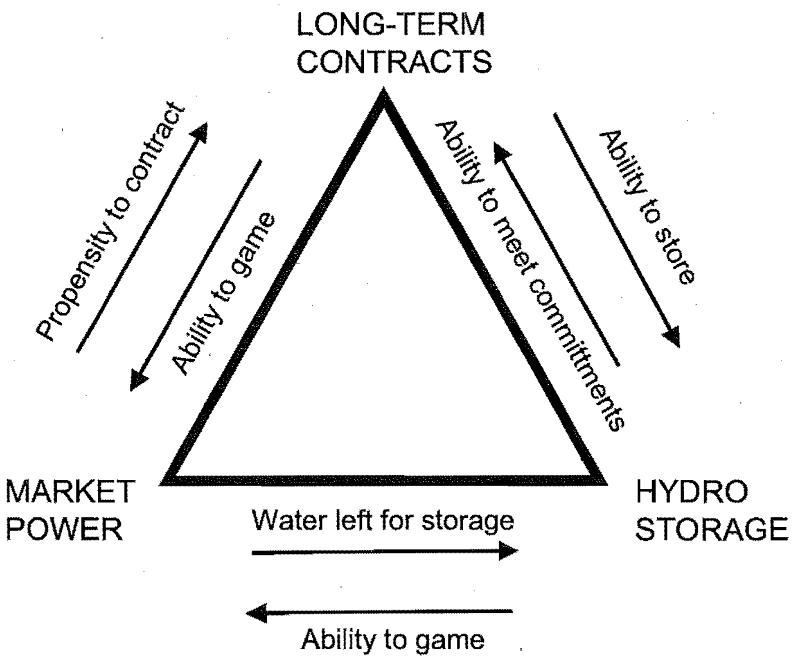


Figure 3.2 Cohesive risk management

It becomes clear that the interactions between the three corners of the risk management triangle require a cohesive analysis. The risk management triangle, which forms the basis of this thesis, proposes a decision making framework that is a synthesis of the individual effects of contracts, market power and storage, and the relationships between them (Figure 3.2).

3.3.5 Managing Risk with Storage, Gaming and Contracts

If we assume that the behaviour of a risk averse decision maker is reflected by a quadratic utility function, her expected-utility maximising decision can also be represented by a mean-variance objective function. Notwithstanding the well-documented criticisms of the quadratic utility function (see Appendix A), we choose to employ this assumption for two reasons. First, it is analytically convenient, since it does not require us to make any assumptions about the distribution of wealth faced by the risk-averse decision maker (Markowitz 1959). This freedom is important to the analysis undertaken in Chapters 10 and 11. Second, it explicitly represents an optimal risk position as the best tradeoff between mean profit and the variance of profit, given the firm's particular degree of risk aversion. This aligns itself naturally with a major hypothesis of this thesis, that firms deliberately increase the variance of risk-averse consumers' electricity costs as a way to increase their own profits. Firms will achieve this position by strategically using each of the (interdependent) elements (on profit mean and variance) in the risk management triangle. Traditionally, market power is seen as a method of increasing profits, and contracts are generally believed to reduce risk. However, the insights developed in this thesis suggests that these roles may be partially or completely reversed, with gaming a potential method of stabilisation (or destabilisation), and contracts a source of significant profits for a dominant firm.

These aspects are examined in more detail in the chapters that follow. Before we proceed, however, we should consider the role of the risk management triangle in the context of who the decision maker actually is.

3.4 Who Manages Risk

Within a firm, risk management can happen at a variety of management levels. For a hydro firm, a variety of risk management techniques may be employed by anyone from well-diversified shareholders to the manager who controls the half-hourly release from the reservoirs. This raises the question of whose responsibility it is to perform risk management using the framework proposed here. Different levels of management and ownership within a company will have different attitudes to risk, and thus the beliefs of what the best tradeoff is will differ. Shareholders who own portfolios of stocks may not be concerned about variability *per se*, but of its covariance with their other investments. Under the CAPM framework for example, higher risk is usually associated with higher return, so well-diversified shareholders may prefer a higher level of risk than management would.

One could imagine decision making within an electricity generation company happening on two levels. On a high level, a strategist is concerned with the long-term profitability and survival of the firm. On an operational level, the tactician receives directives and/or incentives from the strategist and operates the generation assets of the firm accordingly. The information passed down this hierarchy may be storage targets, desired position in the merit order, contract levels or even broad behavioural parameters. Figure 3.3 describes this situation.

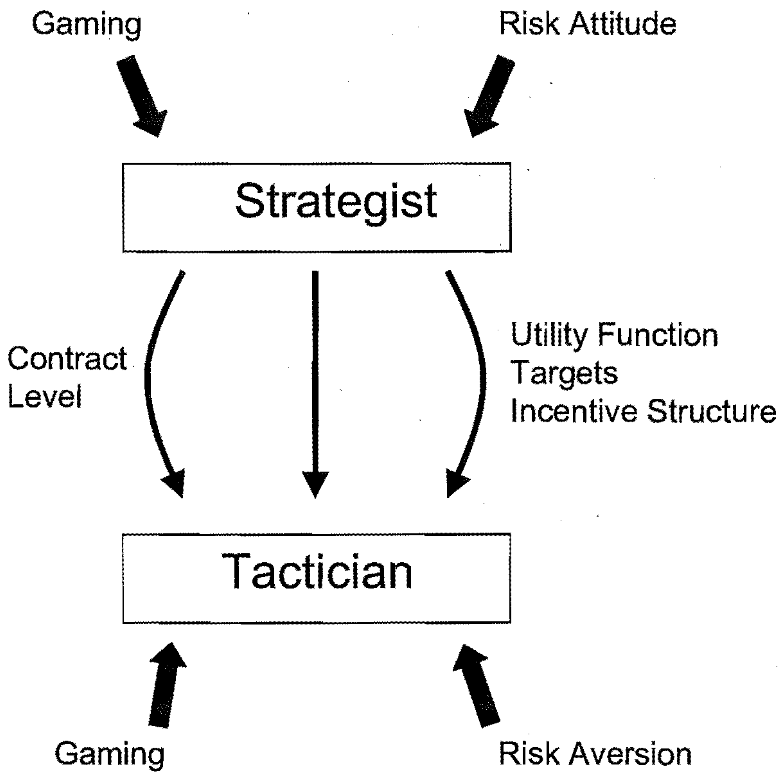


Figure 3.3 Strategist – Tactician information flows

Much of the modelling in the literature implicitly assumes that these parties are one and the same. We traditionally model a single supply firm, simultaneously operating in the spot and contract markets. This is a valid model, if we believe that the decision as to the strategy in the contract market is in the same hands as those controlling spot and reservoir decisions. We also make the implicit assumption that the risk attitude represented by the utility function is common to all three decisions.

Separating out these two roles means we can further enrich the model. Brander and Lewis (1986) described a model which analysed the effect of financial structure (or debt leverage) on output decisions in a Cournot duopoly. They reported different output decisions depending on who was in charge of the operations (tactical side) of the firm – the bondholders or the equity holders (shareholders). Bondholders concern themselves with maximising returns in bad states of the world, where they become residual claimants, whereas shareholders aim to maximise returns in good states of nature, where they are residual claimants.

Both the tactician and the strategist have certain attitudes to risk, can be gaming with other market participants, profit maximising individually and/or responding to specific incentives from above. The tactician may assume the contract level is fixed, and operates according to her utility function. The strategist concerns herself with long-term risk aversion, gaming and profitability, but must also consider how best to communicate these incentives to the tactician to achieve the desired outcome.

In the context of this thesis, managing risk is the responsibility of those individuals who have ultimate strategic control of all the mechanisms that are proposed in the risk management triangle. This is likely to be at a high level of management. It is assumed that their attitude to risk reflects company-wide policy, and can be represented by a utility function incorporating risk aversion.

3.5 Conclusion

It is the intention of this thesis to apply the framework outlined here to the case of an electricity generator who faces quantity risk driven by inflow (or, more generally, input) uncertainty.

The following three chapters examine the role each element in the risk management triangle individually plays in determining the level of price and quantity risk a decision maker is exposed to. Chapter 7 considers the interaction effects between these elements. These four chapters also serve as a survey of the literature, which reveals that few existing models simultaneously consider all three mechanisms in the context of risk.

While the exact chapter outline of the thesis was presented in the introduction to this Thesis, it may be helpful at this point, to state the broad aims of this study. Given the literature review, the rest of the thesis intends to:

1. Examine (numerically) the level of risk that a dominant generation firm, using its market power and storage in a profit-maximising, risk neutral manner, is exposed to.

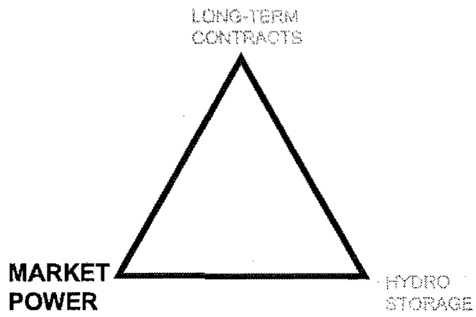
2. Determine how risk averse, mean-variance maximising consumers of electricity will behave in a contract market, given that they are exposed to uncertain spot prices. In particular, we are interested in whether risk premiums emerge as a result of their risk averse, optimal contract decisions.

3. Characterise a long-run equilibrium between contract and spot markets, where dominant generation firms face uncertain input costs, and supply risk averse contract buying consumers. Models will be developed for both the situation where generators do not account for the effect their spot behaviour has on contract prices, and the situation where they do. In particular, we wish to investigate the possibility that these firms can find equilibria that support risk creation, or market destabilisation, in order to increase contract profits.

4. Examine the numerical solutions to the models in (3) for a reasonable set of input parameters.

The final chapter in the thesis (Chapter 13) will draw conclusions and proposes areas for future research.

4



MARKET POWER

4.1 Introduction

Market power could be loosely defined as the ability of firm(s) to alter market outcomes to their advantage (usually to maximise profit), most commonly resulting in prices above the competitive equilibrium, (i.e., where firms set output so that price equals marginal cost).

The use of market power is a major component of the risk management triangle. As will be shown in this, and later, chapters, the ability to influence market outcomes is an important factor in maximising profit and managing risk. This chapter focuses on the former, more traditional, motivation, and provides a review of market power analyses relevant to electricity. It does not account for any combined effects, on risk management, with contracts and/or storage.

Research into market power in electricity markets has occurred on two major fronts. *Ex post* analyses use empirical models along with various theoretical measures of market power to examine whether market power is indeed being exercised in deregulated markets, or at least to determine whether it is detectable. Very recently, available data on market prices has become sufficient to perform credible analyses: this was not possible with the state of the deregulation experiments a few years ago. These analyses are the subject of Section 4.3

Secondly, *ex ante* analyses attempt to provide models that help predict outcomes in imperfect markets where few firms exist. In exerting their influence over market outcomes in a profit maximising manner, these firms must consider their interaction with their competitors, who may or may not be pursuing similar objectives. The vast literature of "game theory" includes models of strategic interaction between firms. Traditional models of game theoretic behaviour have been applied to electricity markets, and further research has been conducted into alternative ways of representing strategic behaviour when empirical observations depart from the predictions of traditional models. Five broad conjectural approaches to oligopolistic gaming are presented in Section 4.4.

The characteristics of different market structures lead to different opportunities and incentives for firms to exercise market power, and the chapter concludes with a discussion of the particular characteristics of small electricity markets, such as New Zealand, that would provide the opportunity for firms to "game" the market. We then consider how appropriate the various theoretical models are to markets such as these. In addition, we will present other general arguments for and against each method of modelling strategic interaction.

4.2 Perspective

Even though market power may only be exercised by the supply firms in New Zealand and other similar electricity markets, it is of interest to all market participants, including regulatory authorities and governing bodies.

Given the above definition of market power, a regulator could detect it by measuring price-marginal cost differentials (which form the basis of market power measures such as the Lerner index). However, a contentious issue is whether the regulator should consider the short-run marginal cost (SRMC) or long-run marginal cost (LRMC) of the firm. A firm that prices above SRMC may not, in fact, be acting anti-competitively, but simply responding to long-run incentives for capital cost recovery. Examples abound in other industries where SRMC pricing is not practised, yet is relatively unquestioned by regulators. Consumers do not often face the true SRMC of long distance telephone calls, or airline seats. In any case, in the context of hydro-thermal electricity markets, the true SRMC of hydro plant is difficult to define.

It is clear that it is difficult for a regulator to assess market power in such a situation. This thesis does not intend to cover any issues faced by a regulator (for example how market power can be detected and inhibited), nor the philosophical view of the social welfare aspects of the use of market power. Market power will be considered from the viewpoint of the firm, and references to marginal cost imply SRMC. Whether the firm restricts output and attains higher prices in order to recover long-run costs, or simply to act in the interests of its shareholders and increase profit, will not be addressed.

4.3 Evidence of Market Power in Electricity

The restructuring of electricity markets has been predicated on the belief that it will lead to competitive outcomes. However, in many countries there remain a few dominant firms and suspicions of anti-competitive pricing (Borenstein, Bushnell and Wolak (2000), Pineau and Murto (1999)). Borenstein and Bushnell (2000) comment that if firms with even modest market shares weren't exercising market power, it would be against the interest of their shareholders. The uncertain future of electricity market regulation almost behoves firms to "take profits while they can" (Borenstein and Bushnell (2000)).

This has led to empirical studies, especially in England and Wales, and in California, that attempt to establish measures of market power, usually based on the disparity between the observed outcomes and the theoretically possible competitive outcome (where participants bid at marginal cost). Others (see, for example, Schmalensee and Golub (1985)) use market concentration as a measure of market power. The majority of these studies have established that market power is indeed being exercised, particularly in high demand periods when a significant proportion of cheaper plant is at capacity (Borenstein, Bushnell and Wolak (2000), Wolak and Patrick (1996), Wolfram (1999)).

The exact fashion in which the market power is exercised is still the subject of much debate, mainly due to the individual characteristics of various markets. In most markets (California, UK, New Zealand, Scandinavia and others) firms submit an offer stack, consisting of a series of prices and the quantity they will supply at each price (which may correspond to different levels of output for individual plant, or different plant altogether). Plant is then dispatched by the central system operator in order of merit, from the cheapest to the most expensive, with the most expensive plant required to satisfy demand setting the market price[14], which is paid to every participant (known as a Uniform Price Auction, or UPA). In markets such as New Zealand, forecasts of prices in each half-hour period of the following day are released daily, which allow generators to review their bids. If a generator is confident that demand could not be met without its generation being accepted, it could increase its bid price, and consequently the resulting market price. The ability of a firm to achieve higher profits with these strategies depends on a number of factors, in particular, firm size, elasticity and magnitude of demand and elasticity of the supply from its rivals.

In low demand periods, the opportunities for market power are less abundant, since demand can be satisfied with cheaper base-load capacity, and the chance of an individual firm not being dispatched at all are greater. In higher demand periods, even a relatively small firm could restrict its offered quantity to force more expensive plant to operate at the margin and bid higher prices. In most markets, the demand side would respond to these scenarios by either storing the product from previous, cheaper, periods, or lowering their requirements (Borenstein and Bushnell (2000)). Unfortunately, neither of these characteristics are evident to any great degree in electricity. Short-run demand (and supply at peak times) is highly inelastic, and electricity is not storable. This results in significant profits for those firms with marginal costs less than price, and an inefficient reallocation of production among price taking firms whose more expensive production acts as a substitute for the cheaper plant being withheld by the dominant firm(s). (Borenstein, Bushnell and Wolak (2000)).

[14] Some markets adjust this price to account for loss of load etc, eg UK

Borenstein and Bushnell (2000) comment that due to short-run demand inelasticity, significant price volatility would be observed as the system approaches total capacity, even in competitive markets. This would be compounded by generation and transmission outages. Mount (2001) cites the use of market power, combined with the UPA structure as the major contributing factors to increased volatility in the Australian market at peak times, even though during off-peak times excess capacity exists.

It can be seen that general consensus exists that opportunities exist in electricity markets for profit maximising firms to affect market outcomes, taking advantage of the special characteristics of electricity markets.

4.4 Modelling Market Power

Economic theory tells us that in a market with a large number of sellers (known as perfect competition), each too small to exert any influence over price, these sellers will set their output levels so that their marginal cost is equal to the market price. The aggregate demand curve for the good in the market is likely to be downward sloping with respect to total output by the sellers, and the price will change as total output changes. However, as these "price takers" do not believe the price will change as a result of any change in their individual output decision, they regard the demand curve for their individual product as horizontal, and will produce the most they can, so long as marginal cost remains equal to or below the market price. We do not intend to discuss the competitive equilibrium in any great depth, other than as a comparison. For a review of competitive models of electricity markets, see Smeers and Boucher (2001).

As discussed in earlier chapters, market power is most commonly observed on the supply side of the market as a result of a small number of large firms. This is particularly true of the electricity market, where the costs of building generation plant are large. In addition, technological factors dictate that plant itself is usually installed in large capacities, making the entry of more firms unattractive as the resulting system total system capacity would be significantly greater than projected demand.

The exact fashion in which profit maximisation is achieved by a firm with market power depends on the number of firms it must compete with. Monopoly outcomes are considered to be the most extreme exercise of market power, where the firm (or a cartel) need only take account of the relationship between its output and demand. However, markets with more supply firms (oligopolies) complicate this, as any change in the firm's output may induce a response by other firms, with the aggregate effect on the market being quite different to that of each individual firm. The nature and magnitude of the response may depend on whether all competitors act strategically or as price-takers, the relative cost functions and capacities of the firms, and what each firm believes about others in the market.

Models of these interactions use game theory to represent how firms will react to each others output decisions, and the equilibrium that results (if any). Five categories of game theoretic approaches used in electricity are chosen for review here:

- **Cournot** competition, where individual firms set output levels in response to their competitors' quantities, assuming they will be held constant,
- **Bertrand** competition, where firms respond to each others' price,
- **Dominant firm** competition, where a "leader" or dominant firm anticipates the quantity reaction of its smaller competitors,
- **Supply function competition**, where firms compete using series of price-quantity pairs, and
- **Simulation** approaches, using variants of the above theoretical models.

Other treatments of market power, including the use of auction theory, will not be dealt with here. In addition, many of the applications of market power studies in electricity are specific to network analyses, where transmission issues affect market outcomes in a unique way. Since such issues are outside the scope of this thesis, only a brief summary of such models will be provided, in Section 4.4.1. The following sections provide a more detailed summary of each approach listed above, and a review of the use of each in electricity models.

4.4.1 Spatial Models of Market Power

The reality of electricity markets is that firms' choice of outputs and/or prices are complicated by transmission issues. Thus, in electricity networks, we observe different prices at different points in the network, reflecting the cost of transmission and the congestion at various points in it. While a firm may not be considered "dominant" with respect to the entire market, transmission capacity may allow it to exercise market power locally. These issues are not accounted for in this thesis, but it is acknowledged that they have an effect on the exercise of market power. The reader is referred to the following papers for various treatments of market power in the spatial setting.

Hobbs (1999) provides two complementarity formulations of a spatial Cournot market. Jing-Yuan and Smeers (1999) also provide a Cournot analysis of a spatial equilibrium. The authors build on earlier work, such as Borenstein and Bushnell (1996), Cardell et al (1996), Schmalensee and Golub (1984) and Oren (1997) (cited in Jing-Yuan and Smeers (1999)).

Hogan (1997) presents a spatial model where a dominant firm competes with a competitive fringe, and later extends this to include a number of dominant firms, competing *á la* Cournot. The dominant firms maximise profit, taking rival offers as fixed and anticipating the response of the competitive fringe. The complicated nature of a spatial model with strategic interactions leads to highly non-convex, non-linear optimisations. Hogan acknowledges this, and presents results with the caveat that they are local optima applying only to well-behaved scenarios found by trial and error.

Traditional dominant firm-competitive fringe theory suggests that the otherwise monopoly (for a single dominant firm) or Cournot (for a few dominant firms) market power of the dominant firm(s) is reduced significantly by the response of the fringe even when dominant firms can anticipate it. Over the portion of the demand curve that the fringe are profitably operating, but not at capacity (i.e., greater than their marginal cost), marginal revenue for the dominant firm(s) is significantly reduced, for an increasing fringe supply curve, and zero for constant fringe marginal costs. Hogan's model reports that transmission constraints allow a dominant firm to "block" fringe production with low prices in some areas, and extract significant profits with high prices in other regions. In this sense transmission constraints allow a form of predatory pricing.

4.4.2 Cournot Competition

Within both the Cournot and Bertrand frameworks, firms assume that their rivals' strategies will not respond to a change in their decision. In the Cournot framework, each firm selects the level of output that maximises profit, given the outputs of its rivals. This could be viewed as each firm finding the monopolistic output based on the residual demand, thus the firm's output is below the competitive quantity. The other Cournot firms perform the same analysis, which raises the question of how equilibrium is found. The Nash equilibrium in a Cournot game is where each firm's optimal quantity is equal to the quantity used in its competitors' optimality condition – i.e., there is no profit-making incentive for any firm to alter its strategy. The equilibrium output can be geometrically interpreted as the intersection of the firms' reaction functions (functions which describe each firm's profit maximising response to its rivals' output). Tirole (1988) provides conditions under which these reaction functions intersect, thus yielding a solution to the game. Linear inverse demand and constant returns to scale yield linear reaction functions that meet Tirole's conditions. Under the same conditions, a Cournot duopoly results in a total output 2/3 of the competitive quantity, or more generally, $(n-1)/n$, where n is the number of firms in a Cournot oligopoly.

The Cournot gaming model has been the subject of much debate in the literature, mainly on the grounds that, in reality, prices are ultimately chosen by the firms, rather than a demand curve. However, its value has been retained because of its computational simplicity, and its relatively sensible solutions[15]. This is especially true when electricity market models include transmission issues (see previous section).

Fishelson (1989) introduced uncertainty into demand and production costs in a Nash-Cournot duopoly. Fishelson then modelled risk aversion with a general class of utility

[15] In that it suggests prices above marginal cost. Compared to Bertrand competition, which suggests that there only needs to be two firms in the market to result in perfect competition, this seems a much more realistic result of the exercise of market power. However, as will be discussed later, many authors consider the Cournot results too extreme

functions, and found that a risk averse firm will respond with a lower output to every level of output of the other firm, i.e., its reaction function was lowered by risk aversion[16]. Note that while total industry output declined under risk aversion, whether an individual firm produced more or less in the Nash equilibrium was dependent on the relative degrees of risk aversion between the firms. This effect was most prevalent in the cost uncertainty model, since uncertainty is firm specific with regard to production cost, but is common to both firms when demand uncertainty is introduced.

To address the long-term implications of uncertain demand growth and capacity investment, Pineau and Murto (1999) developed a multistage Cournot electricity market model, using a special tree structure that "adapts" the Cournot solution to the resolution of uncertainty at each stage. Their paper also provides an excellent review of dynamic gaming models.

Borenstein and Bushnell (1998) provide an extension of the standard Cournot analysis, to the case where the electricity market consists of a price-taking fringe. This introduces the added complexity of flat regions in the residual demand curve, which implies that the Cournot reaction functions have discontinuities. This was dealt with by a numerical grid-search method to find the Cournot-Nash equilibrium. This study was in the context of an empirical analysis of market power in the Californian market, and these authors came to a similar conclusion to Bushnell (2000) (reviewed in Chapter 7), that the potential for the use of market power is greatest in those months in which hydroelectric generation is most constrained by low reservoir levels, and thus residual demand is at its greatest and most inelastic.

The introduction of forward contracting significantly alters these outcomes. A significant number of other electricity models apply the Cournot framework when a forward market is present, but these are reviewed in Chapter 7.

---

[16] Recall that these models do not include the influence of forward contracts. Other studies (see Chapter 7) show that this trend may be reversed when the effect of contracts for supply are considered.

4.4.3 Bertrand Competition

Bertrand competition, like Cournot, leads to fixed-point solutions. Bertrand proposed that instead of competing in quantities, firms undercut their rivals' prices in order to gain market share, as long as price is above marginal cost and the firm has the capacity to meet the demand at any offered price. For two identical firms facing a large number of consumers, a firm can get a share of ½ the industry demand if it charges the same price as its competitor, and the whole demand if it charges marginally less. Bertrand competition differs significantly from Cournot in that it only requires two firms to result in marginal cost pricing, the perfect competition outcome (assuming capacities aren't constraining). This somewhat counter-intuitive result is often referred to as the Bertrand paradox.

However, receiving all of market demand by pricing marginally lower than a rival is, in reality, not always desirable or possible. Resolutions to the paradox are based on a variety of firm and market characteristics which make prices more likely to be above competitive levels with a small number of firms. These include capacity constraints (so that neither firm could actually produce the entire industry demand) and product differentiation Tirole (1988). Another obvious resolution is given when firms exhibit increasing marginal costs – while a firm pricing equal to a competitor may receive half the market demand, pricing marginally below this level may require production at a marginal loss in order to meet the additional demand. Wambach (1999) provided a similar resolution, considering the case of uncertain constant marginal costs. Firms who priced equal to their rivals were now uncertain as to whether their equal share of the market would reap a profit or a loss. If they priced lower than their rivals, they would face n times the profits or losses (for n firms in the market). For risk neutral firms, this did not change the result. However, when firms are risk averse, utility of profit may, in fact, be higher when receiving an equal share of the market than when receiving all of the market. Additionally, since the author conjectures that no risk averse firm will enter the market with potential losses if the expectation of profits are zero, the resulting prices were higher than marginal cost.

Bertrand pricing in electricity markets was initially proposed by Hobbs and Schuler (1985) and by Hobbs (1986), in a spatial setting (see Section 4.4.1). The authors reasoned that Bertrand conjectures were rational for electricity since the good itself cannot be stored, and is thus subject to short-term price competition, although in Hobbs and Schuler (1985) it was argued that the Bertrand equilibrium should be viewed as a lower bound on prices, and limit pricing (where prices are indirectly "capped", by incumbents, at the price that would make the entry of new firms attractive) as an upper bound.

Despite these arguments, very little work has continued to apply Bertrand conjectures to electricity markets.

4.4.4 Dominant Firm Competition

Under this market structure, one firm is believed to have the upper hand (the dominant firm), in the sense that while its competitors' may have a degree of market influence, this level of influence is known by the dominant firm. Optimal output for the dominant firm is determined to be that which maximises profit given the anticipated response of the competing firm(s), which are often either price-taking or Cournot. Hence these models are often referred to as leader-follower models. Stackelberg competition makes the additional assumptions that competition is sequential (i.e., the leader has the chance to make a commitment before the follower) and that competition is in quantities, *á la* Cournot.

A significant proportion of the literature applying variants of dominant firm competition to electricity markets deals with spatial markets, where generation nodes are geographically separated and connected by transmission lines. Many of the results from these papers were specific to network analyses, and are summarised in Section 4.4.1.

Smeers and de Wolf (1997) provide a very interesting case of Stackelberg competition, in which the leader must make its pre-commitment to produce before the resolution of demand uncertainty. The followers, on the other hand, act under complete certainty, both of demand and the leader's choice of output. The model is implemented for the European Gas Market.

4.4.5 Supply Function Competition

Recently, supply function competition has been suggested as a more realistic form of strategic interaction between firms in electricity markets, i.e., that generators submit a range of prices to the market, and the maximum level of output they are willing to supply at that price. A supply function is a series of price-quantity pairs, approximated by continuous or piecewise linear functions mapping price to quantity. Strategic interaction between firms takes place using these supply functions, rather than individual quantities or prices. The seminal work on this form of competition was provided by Klemperer and Meyer (1989). Under somewhat restrictive assumptions, the authors provide a treatment of supply function equilibria under uncertainty, showing that a range of equilibria could be found under the assumption of complete certainty, but that this range is significantly narrowed when uncertainty is introduced. While supply function equilibria (SFE) are generally less extreme than Cournot or Bertrand equilibria, Klemperer and Meyer's supply functions (described by a slope and intercept) can resemble Cournot bids when the function is steep (prices well above marginal cost) and Bertrand bids when the function is flatter.

Bolle (1992) applied the theory to three variants of spot market behaviour (depending on the timing of decisions and recourse) and found equilibrium prices significantly above marginal costs. Green and Newbery (1992) followed by applying Klemperer and Meyer's supply functions to the duopoly in the British electricity spot market. By incorporating certain characteristics of the market (e.g., supply constraints), Green and Newbery were able to narrow down the large range of supply function equilibria. The authors provided an analysis of both symmetric and asymmetric duopoly, finding the asymmetric case resulted in higher prices and profits.

Green (1996) applied Klemperer and Meyer's theoretical framework to analyse the likely impact of divestiture of mid-merit generation capacity in the UK market, and found that removing such capacity from dominant firms would serve to reduce their market power.

Green commented that these results are generally indicative of the market's response to a policy shift (divestiture), rather than an exhaustive analysis of the likely behaviour in the spot market, as the latter will be heavily influenced by the portfolio of contracts held by the generators. The author went on to include long-term contracts in a supply function model in later papers, which are dealt with in Chapter 7.

Both Green and Klemperer and Meyer had modelled the demand side with linear aggregate demand curves representing passive price-taking behaviour by homogenous consumers. Recently Bolle (2001) investigated demand side bidding with strategic demand function behaviour as well. Bolle's analysis was predicated on the fact that many demand-side participants belong to two distinctive groups – retailers or bulk buyers, who act on the behalf of end-users, and industrial users, who are given the opportunity to bid demand as well. Bolle's investigation of the effects of strategic interaction on both sides of the market reversed Klemperer and Meyer's result (that the possibility of large autonomous demand led to marginal cost pricing) to reflect the fact that large industrial users may themselves respond to low prices by increasing their bid quantities, thus returning prices to being higher than marginal cost.

Anderson and Philpott (2002b) analyse the necessary conditions for optimal supply functions that approximate the offer stack submitted by generators. This initial work incorporates a useful representation of the demand uncertainty facing an individual electricity generator. Instead of explicitly representing variation in the assumed demand curve, these authors assume that a "market distribution function" (MDF) is known. This function defines the probability that the firm will not be fully dispatched at any given price-quantity pair in their supply function. This probability is assumed to be a function of both the price and quantity of a particular offer stack, which states implicitly that the market will respond, in some way, to the quantity offered. This representation of the uncertainty gives rise to simple, and relatively intuitive, general optimality conditions on the supply functions. They suggest the use of a non-linear program to find the actual optimal offer stack, but do not consider, in any depth, how the MDF might be constructed. In Anderson and Philpott (2002a), the authors consider the effect on the MDF, and thus the optimality conditions, of assuming that the response of other generators is certain, and demand is uncertain, and assuming that both competitor behaviour and demand is uncertain. Neame, Philpott and Pritchard (2002) made an important simplification to the work of Anderson and Philpott (2002a), by assuming that the market is perfectly competitive and thus is not affected by the individual quantity offered by a generator. Hence the MDF is simplified to just a function of price, and can be formed by making certain assumptions about the distribution of prices in the market. However, none of this work considered the effect of the offer curves, under the variety of assumptions about generator behaviour, on the market equilibrium, which is of critical interest to this thesis.

Recent advances in supply function equilibria have been made in the area of the assumed cost functions. Given the analytical complexity of supply functions, somewhat restrictive assumptions are usually made with respect to marginal costs and demand. While a common criticism of the supply function approach is the potential for multiple equilibria, many authors have noted that the introduction of capacity constraints would restrict the number of possible equilibria (Green and Newbery (1992)). Baldick, Grant and Kahn (2000) presented an supply function approach that allowed the underlying marginal cost function to have a small number of piecewise, continuous segments. These authors proposed a method of "patching together" the supply functions appropriate for each marginal cost segment that allowed supply function equilibria to be found.

Baldick and Hogan (2001) presented a complicated numerical approach to approximating supply functions when the number of piece-wise continuous segments is allowed to be large. These authors note that the sheer complexity of the functions involved, and the difficult characterisation of the equilibrium, means that guaranteeing global profit-maximising supply functions would be almost impossible. However, they do note that again, the presence of constraints (for example, price caps and capacity constraints) will reduce the number of equilibria found.

4.4.6 Simulation Approaches

The use of simulation to represent more complicated (and possibly realistic) strategic conjectures between firms is very attractive. Functions need not be continuous, and solutions do not need to be defined analytically. This allows analysts to include more of the complexities of individual problem situations into a model. However, as with any simulation, the global optimisation of firm behaviour is no longer possible. Hence many authors have incorporated elements of the theoretical analyses, that provide solutions, into simulation models.

For example, Brennan and Melanie (1998) use Cournot conjectures to model strategic behaviour in an empirical model of the Australian power market. "Opportunistic" players examine the effective residual demand for their output (given by predicted load and expected offers by rivals) determine whether opportunities exist for strategic behaviour. If so, they determine monopolistic output over the residual demand, if not, they bid in at marginal cost. While this model does not give analytical results, the authors claim it provides better reflection of the decision processes undertaken by generation firms on a short-term basis.

4.5 Relevance of Theoretical Models to Electricity Spot Markets

It is generally agreed that no individual strand of oligopoly theory accurately reflects the strategic interactions between firms in an electricity market, and that the results from the above models should be viewed as averages or bounds (Hobbs and Schuler (1985). However, it is clear that system characteristics, market rules and industry structure render some models more appropriate than others.

Cournot competition is often criticised in electricity models, since it can yield extreme solutions and requires the definition of price exclusively through the demand function. The market price is the point at which the aggregate quantity offered intersects the demand curve. Electricity demand is highly inelastic in the short-run, and often difficult to specify because of short-term variations, casting doubt on the reliability of Cournot predictions, at least in the short run (Baldick, Grant and Kahn (2000)). Combining Cournot with the short-run perfectly inelastic demand observed in most electricity markets does not yield solutions, and, indeed, using realistic medium-term elasticities can still lead to unrealistic results. Supply function equilibria can be found with perfectly inelastic demand, albeit many of them, from which appropriate solutions must be selected. Supply function analyses are an attractive alternative for those wanting to obtain a realistic view of the way firms, in some markets, actually submit a single offer to the market for an entire day, i.e., using a supply curve. The supply function approach deals with uncertainty by presenting a range of possible operating points (thus keeping the firm operating at "optimality"), whereas the single value suggested by the Cournot and Bertrand frameworks would have to be recalculated in response to each realisation of demand. Supply functions also approximate the situation where a firm has a range of generation capacities available to them at different marginal costs, although at present the discontinuous nature of stepped marginal cost functions makes modelling difficult.

Bertrand assumptions would imply more competitive outcomes than Cournot, as the perceived elasticity of demand is greater for a firm contemplating altering its price than for a firm contemplating an alteration in quantity (Amir and Jin (2001)). As discussed above, Hobbs and Schuler (1985) argued that since electricity is unable to be stored, in the short-run it is subject to price competition, which ultimately results in competitive outcomes. In fact, this was the expectation held regarding how the UK market would respond to electricity deregulation (Wolfram (1999)). The empirical studies of the UK, reviewed in Section 4.2, convincingly report that nothing of the sort occurred. On the other hand, Wolfram (1999) notes that average observed prices are often well below those that the analogous Cournot model would suggest (although studies of European markets provide results that support the Cournot model – see Smeers (1997)). These observations lend further credence to the general supply function results: actual equilibria are on a continuum between these two extremes (Klemperer and Meyer (1989)). Cournot results are observed at peak times, when baseload capacity is exhausted and firms can exercise market power by withholding generation, and during low demand periods, the Bertrand outcome is observed, as firms attempt to ensure their dispatch by bidding in at (close to) marginal cost.

The reality of electricity markets, where prices and quantities are submitted to the market clearing mechanism, appears to make supply functions very attractive. Green and Newbery (1992) and Klemperer and Meyer (1989) cite the fact that supply functions respond optimally to uncertainty in market conditions. However, this advantage tends to be more appropriate in markets where a single supply function must be submitted for a number of sequential periods (e.g., the entire day), and thus must be able to deal with variations in demand over a 24 hour period. However, in markets such as New Zealand and Australia, generators are allowed to change these offers frequently, and often quite close to the period to which they apply[17]. Hence the range of uncertainty that a given supply function must account for, i.e., a single period of up to half an hour, is much less. However, the fact remains that firms compete based on both price and quantity, and thus the supply function approach is at least as good as, and in most situations superior to, Cournot, in reflecting the true nature of the market dynamics that occur.

The attractiveness of Cournot and Bertrand models is in their mathematical simplicity – they lead to fixed-point solutions, as opposed to the "double infinity of equilibria" in the supply function analysis (Newbery (1998)). Additionally, the supply function approach, at present, requires the underlying functions (for example, demand and cost) to be moderately well-behaved. Fehr and Harbord (1993) point out that altering the exact specification of the function can lead to significantly different results, and even the recent advances made with respect to piecewise cost functions and capacity constraints give rise to significant analytical complexity in finding equilibria. Cournot is further supported on the grounds that often a significant quantity of electricity is traded on long-term contracts, so the incentives to undercut one's rivals *á la* Bertrand are significantly reduced (see discussion in Chapter 7 which combines market power and contract effects).

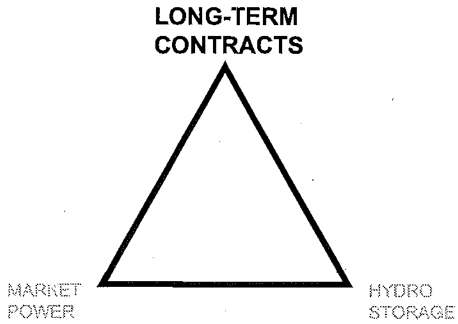
We note that combining Cournot with the short-run inelastic demand observed in most electricity markets does not yield solutions, and, indeed, using realistic medium-term elasticities can still lead to unrealistic results. Supply function equilibria can be found with perfectly inelastic demand, albeit many of them, from which appropriate solutions must be selected. Some have also argued the notion of quantity-responsiveness on the supply side over a single short-term period (e.g., half hour) is infeasible in many markets, especially in peak periods. However, in markets such as New Zealand generators may adjust their bids until two hours before the period itself, allowing at least some strategic

[17] In New Zealand, for example, generators submit separate (and potentially different) offers for each half-hour period, and may revise these offers every half hour up to two hours in advance of the period for which the offer stack applies to. Variation between periods is restricted by ramp rates, which narrow down the range of shifts along a supply function across the course of a day.

reaction[18],. As noted above, though, a firm is likely to make adjustments to both its price and quantity, even if only for one price-quantity "pair" in its offer function.

Notwithstanding the fact that the Cournot approach is largely an approximation to reality in electricity markets, the analytical simplicity of the Cournot approach makes it very attractive to this study, when compared with the more realistic supply function models. For the reasons stated above, we will pursue the current investigation using Cournot conjectures as a model of strategic interaction.

[18] A interesting study to be noted here is contained in Jones (2000), who investigated the correlation (and thus usefulness) of day-ahead prices to final market clearing prices. He found that day ahead prices often bore little resemblance to final prices, implying that a significant period of adjustment occurred over the 24 hour period. However, the duopoly market was dominated by a large firm (ECNZ), which systematically bid a final price of $55/MWh, always greater than, but regardless of what the initial forecast was.



5

CONTRACTS

5.1 Introduction

When a company takes an offsetting financial position to protect itself against fluctuating commodity prices, it is called a "hedge". Hedging is desirable for risk averse individuals since it reduces their exposure to the risk associated with purchasing a commodity. A significant component of financial risk management involves hedging with financial instruments, where the well known option and futures contracts are just the simplest among a plethora of differentiated financial instruments. These financial hedges have their value derived from the spot price of the underlying commodity (hence the term "derivative"), but are often traded on a separate financial spot market.

The traditional motivation for hedging with financial instruments is, as outlined above, to hedge risk. However, in situations where firms hold a degree of market power in the underlying commodity, it is likely they will have a similar degree of market power in the financial market[19] where contracts are traded between sellers and purchasers, giving rise to opportunities to maximise profit with contracts, as well as to hedge risk. This aspect of

[19] In this context, a "market" may be an abstract concept, representing the process by which contracts are bought and sold between parties, rather than a physical location or structure

contract strategy is discussed further in Chapter 7 – here we will restrict our attention to the risk averting motivation.

While the initial waves of deregulation in electricity involved the trading of electricity on a wholesale exchange or spot market, many markets allowed generators to enter into long-term contracts for the supply of electricity with retail companies and industrial firms. Initially, many of these contracts were not as elaborate as in other markets (e.g., oil and agriculture), nor were they as standardised. However, futures and options exchanges are operating in many countries with deregulated electricity markets (for example, NYMEX in the US), although in countries such as New Zealand, the amount of trading that takes place on these financial markets is small compared to the overall level of hedging undertaken by participants. Long-term electricity contracts have largely been financial arrangements between the buyer and seller of the physical commodity, based on simple options and forwards. These contracts are known as one-way and two-way Contracts for Differences (CfD's) respectively. The hesitancy many companies have shown to invest time, money and expertise in developing more elaborate hedging instruments may largely be because electricity derivatives are new, and in many cases deregulation is incomplete.

This chapter provides a description of the two major financial instruments relevant to electricity markets, one-way and two-way CfD's (Section 5.2.1), and will briefly review other types of financial contracts in use in electricity markets (Section 5.2.2). Since many generation firms hold significant levels of CfD's (or broadly equivalent supply obligations), we must understand these instruments fully in order to acknowledge the effect on firm behaviour. However, it does not attempt to provide a comprehensive survey of the vast financial derivatives literature, as many of the instruments available in more developed commodity markets are not currently relevant to electricity, and some never will be. Since the majority of hedging in markets such as New Zealand does not take place on a futures exchange, neither trading strategies nor profit-making speculators (who do not wish to ultimately consume or sell the underlying commodity, but take advantage of spot-forward price differentials) will be included in this review or the analysis that follows it. Rather, it is the intention of this thesis to provide a mid- to long-term assessment of the optimal degree of hedging, in aggregate, that an electricity market participant should engage in.

5.2 Types of Contracts in Electricity

5.2.1 Contracts for Differences

A forward or option contract sets out the following parameters relevant to the physical delivery of a commodity:

- Quantity,
- Strike Price (unit price agreed to pay for the quantity covered by the contract),
- Point of delivery, and
- Time of delivery.

While CfDs are based on options and forwards, which are defined with respect to physical delivery of a commodity, they are defined in terms of the funds that will change hands as a result of an agreement about the above 4 aspects of trade. In this way CfDs are very similar to financial "swaps". CfDs arose because of factors specific to deregulated electricity markets complicating the use of contracts in their "physical" form:

- All producers contribute to a common pool.
- Electricity is not traceable within this pool, hence no producer can guarantee delivery of "their" product to an individual consumer.
- Producers and consumers must sell to/buy from the pool at volatile spot prices, which may never be equal to contract prices.
- "Delivery" is frequent – in some markets every half an hour, or even less.
- "Transport costs" (i.e., transmission costs) are dependent on a very complex set of equations.

CfD's describe the financial transfers between parties "as if" physical delivery as described above had taken place. In essence they are identical to their physical counterparts in that they guarantee a specified quantity at a specified price (the strike price) for a specified time, but also take account of the fact that both parties will trade the physical electricity on a single (spot) market for all electricity, at the prevailing spot price. The contract does not, however, contain any obligation to consume or generate. These aspects are expanded below, for each type of contract. We will begin by examining the more common two-way CfD.

Traditional Forward Contracts

A traditional forward contract describes an agreement between two parties - a seller and a purchaser - guaranteeing physical delivery from the seller to the purchaser at the agreed time and price (the price of transportation is included in the contract). If the spot price is greater than the strike price, the purchaser accrues a payoff equal to the difference, as the high spot prices have been avoided. However, if the spot price falls below the strike price, the seller gains by being able to sell output at greater than the spot price, and thus accrues the spot-strike price differential payoff. These payoffs are conceptual in this regard[20], as they represent costs or revenue avoided. For a seller of a commodity, physical revenue from spot and contract trading can be described as:

$$\Pi = p^s(g-k) + p^f k$$

where g is the total quantity produced by the firm, k is the quantity sold on forward contract, and p^s and p^f are spot and contract prices respectively.

CfD's

Rearranging, we can show that revenue can also be written as:

[20] Unless we account for the purchaser selling back the contracted quantity to the spot market, or the seller initially purchasing it from the spot market

$$\Pi = \left(p^f - p^s\right)k + p^s g$$

This equation shows that should the selling firm receive the spot price for all its output (the second term), as in the case of electricity markets, it should receive or pay the spot-strike differential multiplied by the contract quantity in order to reflect the true payments associated with a physical forward contract.

To achieve this, a CfD guarantees that there will be a financial transfer from the party that benefited from the spot-strike price disparity to the party that was disadvantaged (Figure 5.1). The amount transferred is the first term in equation 5.2, which is necessary to ensure the net payment (i.e., balance of spot and "settling up" amounts) is equal to that which would change hands in the case of a physical forward[21].

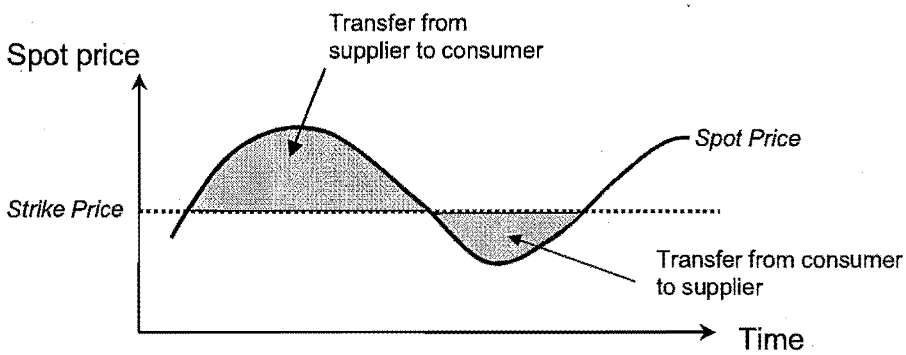


Figure 5.1 Two-way CfD payments

One-Way CfD's (backup contracts)

Conceptually, a call (put) option gives the holder the right, but not the obligation, to buy (sell) electricity off the option seller at the strike price. The alternative for the holder is to buy (sell) their requirements directly off the spot market, at the spot price. A rational decision maker will therefore exercise the option when the spot price exceeds (is below) the strike price. The payoff from the option to such a decision maker is thus zero over the range of prices where it would not be exercised, and the difference between the strike

[21] Two-way CfD's are very similar to commodity swaps in this regard.

price and the spot price for the range of spot prices where the call is exercised. A call could be viewed as a price cap for a holder; a put as a price floor.

In the same way as two-way contracts, one-way CfD's are the financial equivalent of this process. In this case, however, transfer payments are only made in those states where spot prices invoke the option holders to exercise their contracts. In the case of the seller of a call option, equation 5.1 above is rewritten:

$$\Pi = \begin{cases} \left(p^f - p^s\right)k + p^s g, & p^s > p^f \\ p^s g & , \quad p^s \leq p^f \end{cases}$$

In electricity markets, one-way CfD's are often known as "backup contracts". Backup contracts enable a generation firm to purchase electricity from another generator when the spot price exceeds the backup strike price. These contracts can be used when a generation firm faces a plant outage or high marginal costs (preventing it being able to produce electricity profitably), and are especially useful to hydro generators who cannot meet their two-way contract obligations, thus limiting the price risk associated with difference payments at high spot prices (see Section 5.3).

Unlike forward contracts, a price is paid in order to obtain an option contract, and thus one-way CfD's. In addition to the strike price in the contract (which is paid when the option is exercised), and option contract is priced using techniques such as the Black-Scholes formula to reflect the fact that the risk-sharing is not symmetric: while the holder of a call, for example, will benefit from exercising the contract when spot prices are high, the seller of the option will not benefit when prices are low, since the holder will not exercise in these states. A forward contract, on the other hand, has no price *per se*, other than the agreed price of the commodity (although risk premiums may be built into the strike price), since the risk is shared symmetrically between buyer and seller.

Timing Issues

While physical forwards and option contracts refer to transactions that take place at a certain point in time, CfD's cover the exchange of the electricity commodity over a period of time. Given that electricity may be traded every half hour (or 5 minutes in

Australia), it is not practical, or sensible, to have separate contracts defined for every trading period. For this reason, CfD's are sold in "tranches", covering every trading period for a certain length of time. For example, a 1000MW contract for a month entitles the holder to purchase 1000MW of electricity, at the strike price, every period for a month.

For the purposes of this study, we assume this time period is measured in months (usually 12), and that the opportunity to re-negotiate the contract position comes frequently enough to ensure a constant contract cover (i.e., annually for a 12 month contract).

5.2.2 Other Forms of Contracts

The natural extension to a forward contract is a futures contract. Futures markets exist where forward contracts can be standardised to a unit form, and thus bought and sold on financial markets. The major advantage of futures is that if a company wishes to change its hedge position after the initial purchase of contracts (for example due to lower than expected demand), it is a lot easier to sell a standardised futures contract on an exchange than it is to re-negotiate or even sell a customised forward contract. There exists a vast literature on optimising futures trading over time, but for the reasons given in the introduction (i.e., that in many markets, contracts are not standardised to unit form and traded in large quantities), we will ignore futures trading and restrict our attention to forward contracts.

Gedra (1992) defined a contract formed as a derivative of forwards and call options. The "callable forward" was designed as a basis for interruptible load arrangements, where utilities would guarantee consumers either the electricity, or a cash payment to the value of the strike price of the contract. The strike price would be a decreasing function of the probability of interruption, so that consumers who placed a high value on certainty would pay a higher strike price under the contract, but would also be compensated to a greater degree when interruption occurred. Extensions to Gedra's contribution to demand-side management contracts are provided by Oren (2001).

While the majority of instruments used by electricity companies are energy derivatives, Gersten (1999) comments that recently some companies have been investigating weather derivatives, a customised contract that pays when certain weather conditions prevails. Since electricity demand (and often supply) is inextricably tied to weather conditions, these contracts may provide adequate offsetting positions.

5.3 Net Positions in the Market

Since CfD's do not imply any obligation to generate or consume, a certain amount of flexibility accrues to firms with regard to how they deal with their net contracted position. As shown in Section 5.2.1, the transfer payments implied by a contract position are driven only by the spot-contract price differential, thus production and consumption decisions can be considered completely separately[22]. A generation firm geographically[23] separated from another market, for example, is not prevented from selling CfD's in that market. The purchaser of the contract, settles up with the generation firm regarding any spot-contract price disparity, effectively buying electricity off the spot market in which the contract was sold, in order to meet the contractual commitment. This is actually an extreme case of a generator being **over-contracted** in a given market, or contracted for more than they would desire (or be able) to supply to the market (usually in reference to some output benchmark, e.g., the competitive level). Over-contracting refers to the situation where a firm has sold a greater quantity on contract than they can, or choose to, physically supply to the market.

CfD's implicitly allow generation and consumption firms to make physical output decisions that deviate from their contract position in the following ways:

- A generator can be overcontracted, as described above
- A generator can be **undercontracted,** i.e., it generates more than the contract quantity. In this case, the surplus output is sold on the spot market, and earns the spot price.

[22] Although, such an approach would be unrealistic, especially from a profit maximising perspective, as discussed in Chapter 7.
[23] Or at least in terms of an electricity network

- A consumer can be undercontracted, i.e., with additional requirements over and above the contract level, can purchase the extra amount off the spot market at the spot price.
- An over-contracted consumer could also (effectively) sell, on the spot market, any of the contract quantity not used and receive the spot price. This provides efficient price signals for curtailment of load, since if the spot price exceeds the consumer's valuation of the marginal unit of load, he/she will sell it back to the spot market.

These characteristics of CfD's help ensure that correct price signals for marginal consumption[24] and production are maintained in the market, even though a proportion of transactions take place at a fixed price that may be considerably different to the spot price.

5.4 Managing Risk with Contracts

The use of CfD's, or their forward contract equivalent, is well researched, especially in the agricultural sector for hedging wheat and other commodity price risk (Hull (1995)). Forward contracts were mainly used by farmers to sell their harvest forward to avoid the financial consequences of a drop in the grain price. Other than buyers and sellers of physical commodities, arbitrageurs and speculators also trade in forward contracts, mainly to exploit price disparities or to speculate on favourable price movements. Our study here focuses on a problem similar to that of the farmer, namely, sellers and/or purchasers who wish to reduce their uncertainty about the future value of their production and requirement.

Any contract that locks in a fixed price for a fixed quantity removes all uncertainty surrounding the cost of purchasing, or the revenue from selling, the contract quantity. In this way, forward contracts serve to reduce the downsize risk associated with high spot prices for a consumer's load, and low spot prices for a producer's output.

However, this comes at a cost. While suppliers normally benefit from high spot prices, this profit only accrues to the portion of their output not covered by contract. Equally, consumers only enjoy the benefits of low spot prices on any requirements over and above that purchased on contract. For a decision maker that is strictly minimising risk and not concerned about return, this opportunity cost is inconsequential. Such a decision maker can reduce risk by increasing contracting, if no uncertainty about future output or load exists. Ultimately, firms that purchase electricity can completely hedge their electricity input costs by purchasing all future output on contract, and supply firms can completely hedge output revenue risk by selling all future output on contract.

However, in a world without uncertainty, attaining a perfect input hedge for consumers, or perfect output hedge for producers, is not this trivial, and in some cases may not even be possible. We shall address the two sides of the market separately.

Electricity Demand

On the demand side, loads are not perfectly predictable, and for some market participants could be highly variable (for example electricity retail companies, whose customers do not see the spot price, and independently set their load according to a wide range of factors). The possible outcomes are illustrated in Figure 5.2, for a given load probability distribution.

[24] For example, the ability for a consumer to earn profits in high spot-price periods makes curtailment of load attractive.

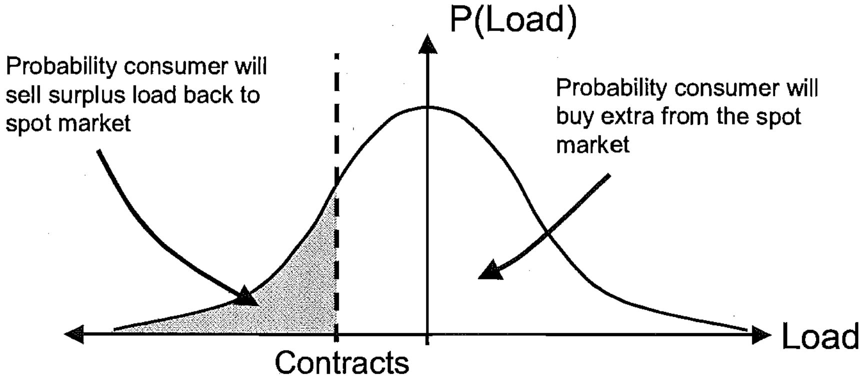


Figure 5.2 Impact of load distribution on net spot purchases

Increasing the level of contracts can be seen to increase the probability that the consumer will have to sell a surplus back to the market, while decreasing the probability that extra electricity will have to be purchased off the spot market. What this represents for the net risk position of the firm is also determined by the distribution of spot price. Under a mean-variance optimisation, total attitude to contracting will be driven by the effect of these two uncertainties on both profit level and profit variance. Consumers' profit levels will benefit from low spot prices if load exceeds contract cover, and high spot prices if the converse is true. Hence the effect of load and price on profit variance is not only driven by their individual distributions but also the covariance between them.

Given the importance of the distributions and covariance of these variables on risk and return in a mean-variance framework, the implications for the best level of contracting for a consumer becomes very complicated, and numerical techniques such as Monte Carlo simulation become attractive to represent the net risk position faced by a firm for a given level of contracting. Alternatively, assumptions can be made about each of the distributions, or their relationships. Chapter 9 investigates a number of these assumptions, and the implications for optimal contracting by consumers of electricity.

Electricity Supply

Not surprisingly, an analogous situation exists for suppliers. Increasing the level of contracts reduces the amount of non-contracted output that will be sold at the spot price, thus reducing the overall variability of revenue and thus risk. However, this is only true in those situations where profit maximising output levels exceed contractual commitments. As outlined in Chapter 2, uncertainty in profit maximising output levels for suppliers, and thus uncertainty about their relative contract position, are driven by two major factors:

- Output risk, caused by plant or transmission outages or breakdowns. For an un-contracted, or low-contracted firm, this may only imply an opportuniy cost in terms of profitable generation that was not possible. However, for a firm that has sold a significant number of contracts, its own generation plant available may not be sufficient to supply the contract amount. The higher the level of contracting, the higher the probability a supplier will be adversely affected by an outage (Gersten (1999)).

- Variability in the marginal cost of generation. Here the firm may still have the **ability** to generate the contract amount, but may have to do so at a marginal loss. Low inflows or high input prices may result in marginal costs being above price.

As outlined in Chapter 2, risk associated with plant outage (output quantity risk) will be ignored in this thesis, although it could be viewed as a limiting case of the second point above, if outage was represented with a very high marginal cost.

Again, increasing contracts may actually increase risk, in terms of financial consequences of having to either effectively purchase the contract shortfall off the spot market at spot prices, or produce the shortfall at a loss. Chapter 8 illustrates a situation in which this is clearly true. This risk can, however, be significantly reduced by holding backup contracts, which gives the generator the option to call on other generation firms to supply any shortfall, should the spot price be high.

In summary, it becomes apparent that two factors determine the mean-variance optimal position for a producer and consumer who hedge with forward contracts: the level of contracts themselves, the distributions of quantity (load or output), the distribution of spot price and the interaction between quantity and price. In reality, the latter three factors may be influenced by the nature of decision making and firm and market structure. As will be shown in Chapter 8, market power influences the price-quantity relationship and may act in a risk-reducing as well as profit maximising fashion, since the level of residual demand elasticity for a dominant firm could lead to a natural hedge (a decrease in the quantity supplied to the market, because of a lack of inflows or higher marginal costs, will be offset by higher prices). Additionally, consumers may reduce load in response to high spot prices, also providing a natural hedging mechanism (Chapter 9).

5.4.1 Strategic Value of Contracts

While the traditional motivation for contracting has been to reduce uncertainty about future wealth, i.e., pure risk hedging, other incentives may act on a firm with respect to contracting. One commonly cited aspect of contracting, applying to use in imperfect markets, is the ability to influence a rival's spot market strategy through the choice of contract level. These issues will be discussed further in Chapter 7.

5.4.2 Contracts for Entrants and Capital Investment

Contracts also provide a degree of certainty about the future utilisation of plant that is not yet built, as pointed out by Newbery (1998), as post-entry prices can be locked in without risk. As discussed in Chapter 2, potential entrants, and capital investment decisions, on the supply side of the industry will only go ahead if their generation capacity is utilised to the extent that, in the long run, the average price received for output covers average cost, including debt servicing and capital recovery (Kriebel and Hornstein (1999)). Selling long term contracts is one way of providing certainty about the long term viability of new plant. Even if average spot prices are below the entry cost, contract prices may provide a return great enough to ensure the investment is profitable (Newbery (1998)).

5.5 Markets for CfD's

A large proportion of the current risk management literature focuses on the economic function of well-developed forwards or futures markets. Because of risk aversion, parties who are trading, or hold a "position", in the physical commodity also purchase futures or forwards as an "insurance" against price changes. Effectively, they desire to transfer the risk of unfavourable price movements to someone who is willing to bear that risk, or to an individual with the opposite risk profile (i.e., both parties are reducing their risk by trading positions). Forward markets facilitate this trade. Two early theories developed the role of futures markets in hedging risk:

- Normal backwardation: Sometimes a hedger wishing to complement a physical position by going "short" (a net commitment to sell) will sell the forward contract to a physical trader who desires to be "long" (a net commitment to buy). Market equilibrium was found when supply (quantity of short hedges) equalled demand (quantity of long hedges). However, as noted by authors as early as Keynes, there are usually more physical hedgers wishing to be short than those wishing to be long. The supply-demand differential was made up by speculators, who received a risk premium (actually a discount to the expected spot price) from the short traders, in order to assume their risk. This disparity between equilibrium futures prices and expected spot prices was known as "backwardation".

- Portfolio theory of hedging: The underlying presumption of portfolio theory is that portfolio owners hold a variety of stocks whose individual risk combine to reduce total risk through the negative covariance of returns. In this way, a trader in a physical commodity "diversifies" by purchasing a forward position in the underlying commodity, whose return offsets the gains or losses made on the spot market. The extent to which traders will purchase forward contracts will depend on their risk aversion, which is assumed to be represented adequately by their utility function (Williams (1986)).

Both these theories have received their criticism (see, for example Williams (1986)). As a result of a large number of arbitrageurs and speculators in well-developed markets, and conversely a relatively small number of participants who wish to see physical delivery of the commodity, these markets result in forward prices that are equal to the best estimate of the future spot price, and provide efficient signals about the future state of supply and demand (Williams (1986)). However, we intend to model a contract market that is not well developed, and to provide a representation of the effect of risk aversion on the desire to contract. Aspects of both theories are helpful in this regard.

Both theories form the basis for a more general supply and demand analysis of a forward market, as outlined by Guthrie (1998). More exactly, while all participants of a financial market "demand" forward contracts, those who wish to obtain a short position (i.e., sellers) can be thought of as suppliers. Contract market equilibrium is found at the forward price where all those wishing to sell short have their contract quantities taken up by those desiring long positions. The extent to which traders wish to hedge with forwards contracts depends on the degree of their risk aversion, as modelled by utility functions of the hedgers' total expected risk and return from the two (financial and physical) positions. Risk averse hedgers can be defined by the following categorisation:

1. Pure speculators, whose utility functions reflect their expected wealth from the spot-forward differential

2. Pure hedgers, whose utility functions represent the variance of their total expected wealth derived from physical commodity transactions, both spot and contract

3. Hedgers whose utility functions are somewhere in between, represented by a mean-variance function of wealth

As with physical commodities, a downward sloping "demand curve for contracts" can be defined by the derivative of the total benefit (utility of total expected profit) function with respect to contracts. Hull (1995) shows that, without loss of generality, the market can be assumed to be completely comprised of the third category. This is because the forward demand equation for these hedgers can be shown to be completely defined by a combination of the forward demand functions of the first two groups.

Expectations

Forward contracts are bought and sold in advance of the period within which the transactions will occur. Since demand curves for forward contracts are derived from the utility of wealth at that future point (or period), they must explicitly consider the hedgers' expectations about the future spot price, future requirements or production, and, in the mean-variance framework, the variance and covariance of these uncertain variables, so that the likelihood of realising the various contract positions relative to load and output (Section 5.3) can be included.

Expectations can be naively formed from past observations, and developed further by methods that place more weight on recent observations, account for seasonality, etc. Expectations can be further enriched by incorporating any knowledge of the variables that affect the future state of the market. For example, hedgers that are aware that hydrological inflows will be lower than average this year may revise spot price expectations upwards to account for this. Asymmetrical information within a market would potentially lead to different expectations about price, making it difficult to aggregate across all hedgers to obtain industry demand for contracts.

As the time period covered by the contract increases, it seems reasonable that the hedger's faith in expectations covering later parts of the contract period will decrease. It could also be conjectured that this would decrease the hedger's desire to be locked in to a fixed price, fixed quantity contract, although this would depend on where the uncertainty lay. If requirements or the level of output were still relatively certain, but the hedger was unsure of price, hedging may still be desirable

It would be ideal to incorporate the effect this reduction in predictive ability has on the "propensity" to contract for longer time periods into the demand for contracting. Very little work has been done in the broad contracting literature on this topic. Some possibilities are discussed in Chapter 9. However, this thesis only covers contracts of single-period length, and assumes negotiations are always taking place for the coming period.

5.5.1 Contract Market Equilibrium

If we assume that price expectations are rational and symmetrical, the equilibrium contract price (the forward price that matches long with short positions) reveals that risk premiums exist when more parties wish to hedge long than short, and risk discounts (forward price less than expected spot price) when the reverse is true (Guthrie (1998)). If risk neutral speculators were allowed to trade in the market, these premiums or discounts would be arbitraged away.

In well-developed markets, contracts are bought and sold on an exchange which operates 24 hours a day. This would introduce extra complexity into a contract analysis, since at any point in time a firm could hold contracts that cover a wide range of periods, and could adjust their contract position continuously.

However, we model a market where contract decisions are made much less frequently, and traders are excluded from the market. This may be seen as a reflection of the complex contract negotiation process between generators and consumers (as argued by Green (1993)). In any case, this study is less concerned with the process by which the contract market finds equilibrium in any particular contract round, than defining the equilibrium and the implied optimal level of contracting for both sides of the market. Variations in hedge cover within that time period will not be considered.

It is worth noting that the motivation to model a contract market without risk-neutral speculators is largely based on the New Zealand experience. The "market" for CfD's for electricity in New Zealand is very illiquid, and few trades are observed. While contract prices are not public information, we expect that they contain significant risk premiums, especially given the price volatility of recent years.

5.6 Conclusion

This chapter outlined aspects of long-term contracting that are relevant to electricity markets. Given the boundaries of this thesis, some of the aspects will not be included in the analysis, some will be approximated, and others will be modelled explicitly.

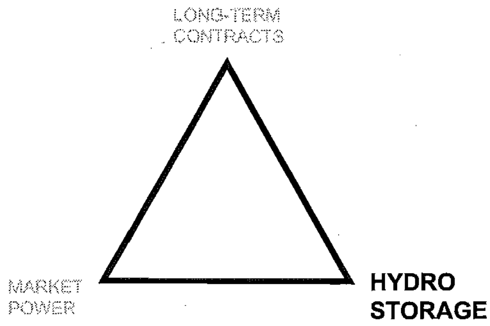
The form of contracts that will be explicitly modelled is the contract for differences, (CfD), a variant of a forward or commodity swap. Option contracts will be ignored, although our assumption that firms are not susceptible to plant outages (see below) could be interpreted as assuming they hold sufficient backup contracts to cover this risk.

Since there exists no significant and well-developed market for electricity forwards in the context of this study, analyses of optimal hedging strategies based on predictions of unbiased futures prices (for example Murtagh and Murtagh (1996)) are irrelevant. The general analysis outlined above is more appropriate since it provides a more accurate reflection of the contract process that actually takes place in the New Zealand industry. A minimal number of speculators exist, and contracts are, by and large, traded between parties that will ultimately consume the electricity, so that asymmetries between risk aversions on each side of the market lead to a contract price different from the best expectation of the spot price. Other more intricate details (such as load interruption, time-of-day variation) will be ignored, so it can be assumed that CfDs are forward contracts that are based on an agreed price and quantity of electricity over a given time frame (usually one year).

This thesis will not provide the best model of forward contracting for hedgers, and will assume that consumers and producers each hold identical, rational expectations about prices. However, we will incorporate the effect of a number of important factors into this expectation forming process. We assume that those who form expectations about uncertain variables such as spot price will make full use of all relevant past information that we can reasonably expect will exist.

The general framework, outlined by Guthrie (1998), for developing functions representing the demand for contracts, will be developed more fully in Chapter 9 to incorporate assumptions regarding the distributions of price and load.

6



STORAGE

6.1 Introduction

A cursory glance over the vast literature relating to storage reveals how important it is to firms and markets. The management of water resources dates back as far as general commodity storage, as water has been stored in reservoirs for irrigation and general consumption for centuries. While general storage principles are relevant here, it is of particular interest to this thesis as to how water storage is optimised in the context of uncertainty or risk. Furthermore, we want to consider the literature that details how a storage manager, in the form of a hydro generation firm, would operate should he or she be risk averse.

This chapter will begin by outlining the general principles established in the commodity storage literature that are relevant to the problem faced by a hydro manager. We will show, though, how the management of water for electricity generation differs substantially from traditional commodity storage management. Finally, we will review the most relevant strand of storage literature – that which relates to the operation of single-reservoir mixed hydro-thermal electricity systems. Here, three methods of modelling reservoir management that hold particular relevance to this thesis will be presented.

6.2 Traditional Commodity Storage Literature

Analyses of optimal storage policies for commodities can be found within the vast storage literature, with major advances being made in dynamic programming techniques by Little (1955). The economic function of storage was initially raised by Keynes who claimed the competitive economy contained insufficient incentives for individuals to store (Keynes (1938)). It is not our intention to provide a full review of the storage literature, as the characteristics of the water storage problem differ substantially from that for most other commodities, as will be explained in the next section. However, the motivations to store remain similar, both from an individual perspective and for the economy as a whole.

The dominant motivation underlying most uses of stockpiles is that it is desirable to hold some level of output back from immediate use as a buffer against uncertainty. In inventory models, stocks can be held as protection against uncertain demand. In agricultural settings, stockpiles of farm output are built up in times of surplus harvests, and, depending on the 'decay' of the commodity, can be spread out over a number of future periods, some of which may be times of shortage. It becomes clear that storage thus plays a vital role in the stabilisation of commodity markets – storage acts as a buffer against the effect of season and climate on agricultural prices. However, as pointed out by Williams and Wright (1991), this effect of storage on price is notably asymmetric:

> *"Because, collectively, the market can always store whereas it cannot borrow from the future, storage is much more effective at supporting what would otherwise be very low prices than at reducing what would otherwise be very high prices" (Williams and Wright (1991), p2)*

In markets where the price for a commodity is driven by the balance of supply and demand, this intertemporal spreading of a commodity tends to produce serial correlations in price over time. A price in a given period is not independent of the market equilibrium in previous periods – if market supply can be supported with production from previous periods, an obvious dependence arises, at least while the storage is available. Even in periods where storage has been exhausted, the incentives to hold back supply from the market may be noticeably influenced by the lack of carryover from the previous period.

This short-run inter-temporal effect of storage is one of many that have often been overlooked in traditional economics, where long-run analyses look for comparative statics, or argue that the sole purpose of storage is the removal of uncertainty (Williams and Wright (1991)).

The motivation to store often goes beyond market stabilisation, especially when storage is in the hands of private owners. Stocks procured at a low price (in times of surplus) can be held back from the market, and sold at times of high prices. Private storers thus hope for periods of bad weather and shortage, and because of product deterioration may prefer a more variable market where times of surplus are followed quickly by times of drought.

Regardless of the motivation, the concept of trading between periods in this manner becomes a central theme in any analysis of storage. Williams and Wright (1991) comment that the storage problem is made interesting by the interaction of uncertainty, non-negative storage and the role of time in production, aspects common to both water and commodity storage. Models of storage, by and large, attempt to find the optimal decision, in terms of the amount to reomove from, or add to, storage, in response to a certain scenario (usually a realisation of uncertain input and/or output prices and/or quantity). This relationship is called a storage rule, and makes techniques such as dynamic programming, in particular Markov decision processes, appropriate as solutions methods.

However, the characteristics of water render many of the assumptions of these traditional analyses irrelevant, and we now turn our attention to models that focus on the hydro-electric power problem.

6.3 Electricity and Water Resources

Water differs from normal storable commodities mainly in that it is relatively costless to store. Storage costs in the traditional sense are usually defined in terms of alternative uses of limited space, the financial cost of having funds tied up in stored commodities paid for but not used productively, handling and transportation costs, and deterioration of stock over time. Where different commodities compete with one another for storage space inside a facility, "capacity expansion" of storage facilities for an individual commodity often refers to the additional allocation of existing space (and consequential reduction in capacity for the competing commodity). In this sense, the marginal cost of expansion is relatively well defined (Williams and Wright (1991)).

The storage of water differs from these analyses in that:

- Reservoirs are usually built exclusively for the purpose of water storage, and once the reservoir systems are built, water transportation and storage is costless.
- As a consequence, capacity expansion involves significant fixed costs, meaning the marginal cost of expansion is difficult to define.
- No price is paid to procure water (although the development of water markets where it attracts a price challenges this assumption).
- The water does not deteriorate over time, although wastage could occur through evaporation and seepage.
- The productive use of the marginal unit of water increases with the amount in storage (head effects).

Hence the marginal cost of storage is close to zero until storage capacity is reached, at which time it increases sharply, as the marginal unit of water requires additional reservoir capacity in order to be stored. However, water that is used for a productive purpose, such as electricity generation, must still be priced to ensure its optimal use. Historically, water in electric power systems has been priced to reflect (i) storage capacities and (ii) the opportunity cost of using water to avoid more expensive thermal generation. As Read (1984) states:

> *"....all of the long-term reservoir scheduling models proposed in the literature, deterministic and stochastic, can be seen as more or less sophisticated ways of estimating [the water value] for one or more reservoirs. This water value is the only aspect of the future operation of the reservoir which the decision maker needs to know in order to produce an optimal solution.... "Read (1984) p4*

The development of the water value concept for hydroelectric systems is continued in Section 6.4.2 below.

6.4 Modelling Approaches to Hydro Reservoir Management

A survey of the literature reveals four broad categories of hydro reservoir management models:

1. Simulation Approaches.
2. Heuristic methods.
3. Mathematical programming (MP) approaches.
4. Dynamic programming (DP) approaches.

The merits of each approach are largely driven by their ability to incorporate uncertainty, and handle the inherent complexity of hydroelectric systems. This complexity includes head effects, dependencies between multiple reservoirs on a single river system, and environmental and operational constraints on release from the reservoirs.

We will discuss the importance of inflow uncertainty to reservoir operation in Section 6.4.1. Other than in the short term or real-time decision making context, where much of this uncertainty is resolved, it seems unlikely that a deterministic model, whether an MP or DP, would provide anything of use to an analysis of decision making under uncertainty or risk (as is the case with inflows, let alone market behaviour). However, we do briefly review a deterministic decomposition model below, as it provides a useful background to the development of the concept of the optimal storage policy. In particular, the notion of a marginal water value, of great importance to the stochastic models that attract a greater focus here, is introduced here.

We will not consider simulation and heuristic models. As argued by Yang (1995), simulation is restricted to a limited set of decision alternatives. Even though it is able to evaluate those decision rules over a wide variety of inflow sequences, thus giving the analyst a comprehensive assessment of the performance of the rule under uncertainty, simulation does not optimise[25]. Similarly, heuristics, while being able to cope with significant complexity, do not guarantee the optimal solution. It is the purpose of this thesis to integrate hydro management with other tools of risk management (market power and long-term contracts), so a model that provides an optimal policy over all decision alternatives is required.

Stochastic programming (SP) models applied to reservoir management, such as Scenario Optimisation and Bender's Decomposition, are generally more complex and difficult to solve than deterministic models (Yang (1995)). Both types of SP models have their treatment of uncertainty limited by the number of scenarios accounted for, and, since probability distributions are assumed to not change as a result of decisions, they cannot incorporate the effects of market power, an aspect that is important to later discussion. We will not consider these models any further.

It should come as no surprise that dynamic programming applications to hydroelectric reservoir management abound. Where mathematical programming applications struggle with the size and complexity of water resource systems, dynamic programming is able to effectively decompose a problem with a large number of (often nonlinear) variables in a number of subproblems, and solve them recursively (Yeh (1985)). The "cost" of this decomposition is that each subproblem must carry a state variable representing the connection to the rest of the optimisation. As the system becomes more complex (e.g., multiple reservoirs, inflow uncertainty), and more state variables are required, the so-called "curse of dimensionality" increases the computational requirements for a DP solution.

DP approaches to reservoir management have been both deterministic and stochastic. The reader is referred to Yeh (1985) and Yang (1995) for a more detailed review of the general development of dynamic programming and its application to reservoir problems. Deterministic applications will be ignored for the same reasons as given above. A review of one important strand of stochastic DP, namely Dual Dynamic Programming (DDP), is

[25] Although increasingly, as noted by Yeh (1985), some degree of optimisation is being embedded into simulation models. However, these optimisations tend to be single period decisions where the number of dimensions are small, rather than a medium-long term optimal storage or release policy.

presented in Section 6.4.3. The use of DDP has somewhat alleviated the curse of dimensionality by using the dual prices on many of the physical (primal) constraints as the state variables, thus allowing a more direct specification of the optimal policy.

Finally, Section 6.4.4 will investigate an important extension of dynamic programming incorporating uncertainty; that is, where the decision maker is risk averse.

6.4.1 Inflow Uncertainty and Risk

The key departure of stochastic reservoir models from their deterministic equivalents is that inflows are assumed to belong to a stochastic sequence. Deterministic models tend to assume inflows will behave according to a particular observed historical sequence, or the expected value of a known probability distribution. While evaluating model performance for a given value of an uncertain parameter can be useful, explicitly incorporating uncertainty into the decision making process is an important extension:

> *"Although the uncertainties of some parameters may be taken into account or be alleviated through a sensitivity analysis, however, the procedure does not explicitly consider these uncertainties and may not lead to satisfactory results." Yeh (1985), p 1799*

The fact that the majority of inflow uncertainty is resolved at the time a release is made does not remove the necessity for modelling uncertain inflows. In many cases, not only must the release decision be made in advance of the period itself, the ability to respond in real-time to an observed inflow sequence may be restricted through the plethora of release and/or streamflow constraints that operate on hydroelectric systems.

The obvious way in which to introduce inflow uncertainty into the optimisation is by way of a probability distribution. The objective of the problem then becomes to maximise the expected value of the benefits from a given release policy, or a variation on this basic idea. Thus a distribution of uncertain inflows gives rise to a distribution of uncertain profits for the hydro firm within the period. This would enable a model to incorporate a utility function to represent risk averse decision making behaviour (see Section 6.4.4).

However, it is the way in which the specification of the probability distribution of inflows is made – more particularly, the statistical assumptions made about its behaviour – that is a source of controversy in the literature (Yakowitz (1982)). A significant number of reservoir management algorithms assume independence between successive inflow realisations, i.e., for an observed inflow i_t:

$$p(i_t \mid i_{t-1},\ldots,i_1) = p(i_t) \tag{1}$$

Such an assumption gives rise to "more elegant and easily computed solution strategies" (Yakowitz (1982), p685). However, the earliest treatment of stochastic dynamic programming and reservoir systems, that of Little (1955), argued that such an belief was unrealistic. The more general Markov treatment of Little describes the inflow process as serially correlated:

$$p(i_t \mid i_{t-1},\ldots,i_1) = p(i_t \mid i_{t-1}) \tag{2}$$

While Little did not statistically support his assertion of correlated inflows, subsequent studies (e.g., Yakowitz (1973)) have shown that this belief was not without sound basis, although in certain hydro systems the correlation was more likely to be on a daily basis rather than the two-week decision interval chosen by Little. One could speculate, however, that with the advent of more recent research into longer-term weather patterns (for example, the El Nino Southern Oscillation, ENSO), correlations over weeks and months could have a more noticeable effect on the shape of the probability distribution function for hydrological inflows. This is of critical importance to hydro systems with small reservoir capacities (relative to demand) such as New Zealand, as the realisation of a number of low inflow-weeks may indicate the beginning of a drought period, which should have specific implications for reservoir management policy, particularly under risk aversion. An independence or "lack of memory" assumption would presume that a high inflow-week, negating the effects of a low sequence, would be just as likely as a continuation of the sequence. Yang (1995) provided a DDP model of a mixed hydro-thermal system which accounted for correlated inflows, and yielded improvements (in terms of reductions in thermal costs) of around 5%.

6.4.2 Deterministic Models

Read (1984) reported a deterministic decomposition model that applied economic principles, developed earlier for the general water resources context, to the scheduling problem faced by the manager of a centrally co-ordinated power system. Central to these analyses is the concept of a "natural value of water", or "water value". Since the residual load, after hydro release is scheduled, must be met from thermal generation, the marginal cost of the most expensive thermal dispatched defines the marginal value of water at that time. In the absence of storage capacities, reservoir managers will "trade" water between periods, in a manner analogous to economic arbitrage, so that the marginal value of releasing water is equalised across periods (ignoring release and storage constraints). If this is not the case, savings in total system cost could be achieved by releasing more in the periods where water is more valuable (due to avoidance of higher generation costs), and less in those periods where it is not. Read reports important general principles characterising storage policies:

1. Water should be released if and only if the marginal benefit from doing so exceeds its marginal value if stored for the future.

2. In optimality, the marginal value of releasing the water in the current period should equal the marginal benefit of storing it for a future period.

The "marginal value of storage" is obtained using the same principles as the marginal benefit of the use of water in the current trading period, i.e., the thermal costs offset, although it may be discounted. Read goes on to incorporate release and reservoir capacity constraints into the decomposition framework. These bounds on 'primal' variables naturally result in intuitive bounds on the dual variables. The marginal value of water when the reservoir is already full is zero.

6.4.3 Dynamic Programming Models with Uncertainty

Read (1989) built on earlier stochastic dynamic programming models (where the state space was the level of storage) and produced a variant of SDP which used a discretised state space based on the price of water. Such an approach was termed Stochastic Dual Dynamic Programming (SDDP). Periera and Pinto (1991) independently developed a modelling technique for reservoir optimisation, to which they gave the same name. Pereira et al used a combination of simulation using historical inflow sequences and stochastic optimisation which concentrates on the range of storage scenarios through which the simulation passed, to give upper and lower bounds on a cost function. At each decision period, the method produces a piecewise linear approximation of the future cost function (which provides an approximation of the marginal water value function). Rather than produce exact decision rules for the entire state-space and planning horizon, it focuses on producing a good solution for the first period, and only forms approximate decision rules for later periods to the extent that this seems likely to significantly improve the initial decision for a small set of inflow scenarios.

In the context of Read (1989), SDDP accurately reflected the decision making process in a centrally co-ordinated hydro thermal scheduling problem and produced significant gains in accuracy and efficiency. Optimal release decisions were defined by setting the marginal value of releasing water equal to the marginal value of storing it for later use. The marginal value of storage was represented by the end-of-period marginal water value curve, while the marginal value of release was defined by the stepped curve representing the marginal costs of the thermal stations that would be incrementally required, as release dropped. The two curves were added, with the resulting curve resembling the original marginal storage curve except for "flats" inserted at each thermal station's marginal cost. This reflected the nature of the dispatch process, where a thermal station would be base-loaded if its marginal cost was less than the marginal water value. The curve was then adjusted for the uncertainty of inflows to produce the beginning of period demand curve for storage. The model was solved by backwards recursion, as in traditional dynamic programming, to find optimal storage policies.

The SDDP method of reservoir modelling was further developed by Scott (1998) to include long-term contracts and participant gaming, and is discussed in Chapter 7.

6.4.4 Models with Risk Averse Reservoir Managers

By implication, the stochastic models reviewed above assume that reservoir managers behave in a risk-neutral fashion, since the optimisations maximise (minimise) the expected value of profits (costs). Given the place that uncertainty holds in the vast majority of storage models, it seems surprising that few of them have considered the role of risk aversion in finding optimal storage rules. Williams and Wright (1991) claim this would complicate the analysis, as many of the effects introduced by storage could be misattributed to risk aversion, rather than a consequence of a collective response of the market. Additionally, the authors claim that due to the serial correlation introduced by storage, the representation of risk aversion becomes critical, especially whether it is measured as risk aversion over income within a period, or over wealth over time:

> *"Are producers more averse to a bad year if it follows a string of bad years?...If the answer is yes, producers' utility functions are not additively separable, an implicit requirement of most conventional analyses based on expected utility." Williams and Wright (1991), p14*

Despite this scepticism about the ability to inorporate risk aversion into storage, Kerr, Read and Kaye (1997) developed a "Stochastic Utility Maximising Dynamic Programming" (SUMDP) model of hydro storage for a mixed hydro-thermal firm operating in a competitive market with uncertain inflows. This model built on the earlier work of Ranatunga (1995), who modelled the thermal commitment problem where uncertainty was in price, rather than inflows. The objective function is to maximise expected utility which was defined over end of horizon cost (wealth) and storage. The non-separability of a non-linear utility of wealth was handled by allowing the state vector to include a running total of accumulated benefits.

Simulations of (a) different degrees of risk aversion towards wealth and (b) different annual inflow sequences yielded the following results. While a risk-neutral decision maker (i.e., strict cost minimiser) would respond to high demand periods by increasing release (to avoid higher use of expensive thermal stations), risk aversion caused the incentives for avoiding low-wealth (i.e., high cost) outcomes to grow, while the associated incentives to achieve high wealth (low cost) became weaker. Thus it became more acceptable for the risk averse reservoir manager to incur larger thermal costs when the uncertainty was such that more water in storage was desirable as a way to hedge against bad outcomes in the future. Risk aversion caused less water to be released, and thus more to be stored. The resulting distribution of wealth shrunk, as very high- and low-wealth outcomes occurred with lower probability. Interestingly, this effect was more pronounced for good outcomes.

6.5 Conclusions

In summary, we have found the following insights, from both the general commodity storage literature and more specific hydroelectric storage models, that are relevant for the optimal operation of hydro reservoirs:

- With the use of reservoirs, water can be traded between periods to reduce the effect of variable inflows. This trade can only take place in one direction, i.e., water can only be stored for later use, within the upper limit on the reservoir. However, due to lower bounds of water storage impose a strict limit on the extent to which water can be "borrowed from the future" for earlier use.

- In this way, storage "smoothes" the normally variable pattern of prices over time, as the commodity is held back from the market in times of surplus and used in periods of shortage. This results in serial correlations in market price.

- Ignoring release constraints and reservoir capacity, the marginal value of releasing a unit of water in a mixed hydro-thermal electricity system can be indirectly defined as the marginal cost of the station that would be called on to generate should that unit not be released.

- The marginal value of holding a unit of water in storage can be defined (by backwards recursion) as the expected value of releasing it in a future period (accounting for storage constraints).

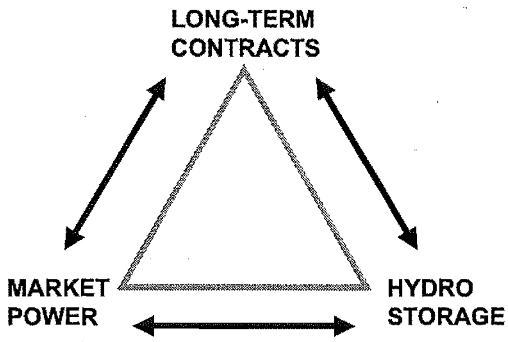
- To reach optimality under risk-neutrality, water should be traded between periods so that these marginal values period are equal in each period, at least as far as storage capabilities permit. More exactly, in each period, release should occur until the marginal value of releasing an additional unit of water is equal to the expected marginal value of storing it for later use.

- Risk aversion causes hydro managers to release less, and thus store more than under risk neutrality

The general DDP framework provided by Read (1989) and later developed by Scott (1998) incorporates all these insights, except risk aversion, into hydro management with uncertain inflows. One area of concern with Scott's model is the absence of correlated inflows. As identified in Section 6.4.1, this is of reasonable importance for a tightly storage-constrained system such as in New Zealand, which largely motivated Yang (1995) to incorporate correlated inflows into a DDP of reservoir management. However, as will be discussed in Chapter 7, the aspect of Scott's model which is important to this thesis, is the presence of gaming between market participants, something absent in Yang's work. We recommend that the future development of risk-averse reservoir operation, especially in countries such as New Zealand, incorporates inflow correlation..

Scott's model, which features gaming between market participants (and is thus discussed in more depth in Chapter 7), will be used to determine the level of variability in profit for the hydro manager that remains after water has been traded between periods in an optimal sense (equalising marginal water values), and this analysis is provided in Chapter 8. The results provide an estimate of the residual profit risk without allowing for risk aversion, i.e., how effective the optimal use of storage capabilities is in providing a buffer against inflow variation, and thus profit risk, faced by the firm. Once optimal reservoir management is aligned with gaming and contracting strategies, risk aversion can then be accounted for.

We now turn our attention to reviewing that model, and a number of other models which have providing a level of synthesis between the three mechanisms of the risk management triangle.



7 INTEGRATED ANALYSES

7.1 Introduction

In the previous three chapters, we have outlined aspects of the three risk management mechanisms important to this thesis; namely contracts, market power and storage. In this chapter, we will show that there are few analyses that provide a synthesis of all three mechanisms and their effect on output decisions of the firm. Even fewer exist that look at the implications for risk management.

There exists some literature, though, that investigates any two of the three mechanisms simultaneously, which can be categorised by the side of the triangle it represents:

(a) Analyses of Storage and Market Power

(b) Analyses of Storage and Contracts

(c) Analyses of Contracts and Market Power

As the next chapter will show, the residual profit risk observed in reservoir management models where the firm possesses market power is relatively satisfactory, even for a risk-averse individual, at least for the system we have studied. Thus, the importance of storage in the management of risk is somewhat diminished. We can deduce, then, that forward contracts and market power become the key risk management tools, within the framework presented in Chapter 3. This has two consequences for the literature review in this chapter:

- Since (a) and (b) consider storage, along with either contracts or market power, they do not warrant a great deal of attention. A summary of a few examples from the literature is provided in Section 7.2 and Section 7.3 respectively. Storage will, however, be considered in depth in the review of papers that include all three corners of the risk management triangle (Section 7.5).

- Extra attention will be paid to (c), especially those papers that consider the aspects of gaming behaviour (Cournot) and contract type (CfD) selected as being relevant to this thesis in Chapters 4 and 5. Relevant papers are reviewed in Section 7.4.

This chapter concludes by outlining the setting within which this thesis will make its contribution.

7.2 Storage and Market Power

Newbery (1984) motivates his analysis by suggesting that the market stabilisation role played by storage in competitive markets (see Chapter 6, Williams and Wright (1991) in particular) is even stronger for firms with market power. He develops storage rules for dominant producers that, instead of arbitraging expected returns to current and future price, trade off current marginal revenue with future expected marginal revenue (in the face of uncertain output). The results depend significantly on the elasticity of price, but in general, Newbery shows that dominant producers prefer stable prices: whether they achieve this through altering their output decisions or increasing the aggregate level of storage depends on the response of price to a change in supply, relative to marginal cost. In the case of linear demand, for example, storage (and spot price) is significantly higher.

Dalziel (1987) provided an analysis of the problem faced by a firm who stored and priced water for further use (such as irrigation). His optimal-control model was closely aligned to analyses in the exhaustible resource literature, and he considered both a monopolist's and a perfect competitor's actions in developing optimal storage rules. Storage policies were derived implicitly from the price charged for water by the firm, and, in particular, Dalziel's model showed how this price path over time deviated from the "natural" price of water, the latter derived from the market price resulting from a release equal to inflow (or no-storage) policy. Dalziel's storage 'period' was a finite-length continuous time interval that contained a high-inflow and low-inflow season (in that order, chronologically). The storage rule developed, not unlike Hotelling's Rule, was that a monopolist (perfect competitor) would set water prices, and thus derive an optimal storage pattern, such that the marginal revenue (price) received from selling water increased through time at exactly the discount rate. This gave two critical points within the time period: when the firm would begin storing water, and when stored water would run out. Within these points the firm would hold back water from release, and thus the price would deviate from the 'natural' path. Reinforcing the results discussed in Chapter 6, both competitive and monopoly firms would price water higher in the earlier time of surplus, and release it in the time of shortage, thus pushing prices below the run-of-river equivalent in the later period.

Dalziel's key conclusions were that the monopolist began storage earlier, and thus stored more water (in the linear demand case) than the perfect competitor, while storing the same amount as a perfect competitor in the constant elasticity case. However, the lack of storage bounds was a significant omission from Dalziel's work, as was the lack of capacities on release. Firstly, the extent to which a monopolist can exceed the competitive storage policy must surely be limited by an upper bound on the reservoir – a firm cannot withhold water from the market indefinitely. Secondly, a critical aspect of Dalziel's work was the connection between the behaviour of the demand for water, and the demand for the use of the water (e.g., irrigation, or electricity production). Dalziel assumed these demand functions were identical. However, while the demand for the product may exhibit the type of continuous behaviour described by a constant elasticity curve, the demand for water cannot, as it will be limited by bounds on release.

Bushnell (2000) used a mixed complementarity problem (MCP) formulation to represent firms with market power (following Cournot conjectures) facing a competitive, thermal fringe. Each of the Cournot firms had a mix of hydro and thermal or nuclear generation capacity. Reservoir management policies were derived directly through the use of water values implied by the dual variable on the energy balance constraint, rather than through recursion in a DDP framework (see Section 7.5). Bushnell showed that hydro managers can extract significant profits from the market in the opposite fashion to that proposed by Williams and Wright (1991) – by withholding water from the market in periods of high demand, when competitors are capacity constrained, and storing it to help increase release in the off-peak periods, a major reversal from traditional optimal hydro scheduling practices (for example Read (1984)).

7.3 Storage and Contracts

Authors such as Turnovsky (1983), Sarris (1984) and Hirschleifer (1989), extend this analysis by questioning whether futures markets would exist in the presence of storage under perfect competition assumptions, and, if so, what role they would play in the stabilisation role. These studies include aspects such as separating physical commodity producers and storers, and speculators in the futures market, while ignoring the possibility that physical commodity consumers may wish to take a long position in contracts. Coupled with assumptions of positive and increasing cost of storage, these analyses are well outside the scope of this study.

Fleten and Wallace (1998) proposed a multi-period stochastic programming approach to solving the problem of a hydro manager who has the option of purchasing various types of over-the-counter contracts. The authors proposed that water in storage and contracts can be viewed as assets that attracted returns (that were uncertain due to variable inflows and variable futures and spot prices), and thus portfolio optimisation techniques could be used to solve for the optimal values of the assets over time, while minimising risk. The basic reservoir management model develops a tree structure representing a number of inflow and price scenarios corresponding to each period in the future, which is extended to include futures contracts. Speculative purchases of futures are ignored (futures price is equal to expected spot price), and risk aversion is represented by penalising the amount profit falls short of a pre-specified target in key decision periods. A numerical example shows that for a certain set of assumptions, the risk averse contracting model reduces risk by around 70% (compared to risk neutral without contracts), while only sacrificing 0.7%

of expected profit. The authors briefly comment that the use of futures contracts appear to increase the average release by approximately 8%.

7.4 Market Power and Contracts

The investigation of the effect of contracts on market power (and vice versa) has been performed both empirically and analytically. We will focus firstly on analytical models of market power and contract strategy, and then present three significant pieces of empirical work for electricity markets. As will be seen, these empirical results tend to support the theoretical analyses.

7.4.1 Theoretical Analyses

Various authors have addressed the problem of what effect long term contracts have on firms' ability to exercise market power. Given the attention that deregulated electricity markets have received recently, much of this theoretical development has taken place with the characteristics of the electricity commodity (particularly its non-storability) in mind. A critical categorisation of the existing work, for the purposes of this study, separates studies that analyse the effect of a fixed level of contracts on a firm's influence over spot outcomes, from those which jointly consider optimal spot market behaviour and the optimal quantity of contracts sold by the firm. Both strands of research have been undertaken under a variety of assumptions with respect to the nature of cost functions, strategic behaviour, number of participants, and demand response. However, the general results are summarised in the following two sections.

7.4.1.1 Fixed Contract Quantities

As discussed in Chapter 4, firms with market power maximise profit by setting marginal revenue, rather than price, equal to marginal cost. Figure 7.1 provides a particular representation of this process, for a monopolist with increasing marginal costs. A reduction in quantity from the competitive level will induce a price increase. So, while a firm sells less in the market, the profit lost on those units (area B) will initially be outweighed by the gains from the price increase on those units still sold in the market (area A). As the firm reduces its quantity from the competitive solution, the profit gained initially increases at a much greater rate than the profit lost. The firm continues to reduce its quantity until the difference between these areas is maximised, thus maximising the monopolist's profit (equivalent to equating marginal cost with marginal revenue). The exact point at which this occurs depends on the price-responsiveness (elasticity) of demand, and the underlying marginal cost structure of the firm.

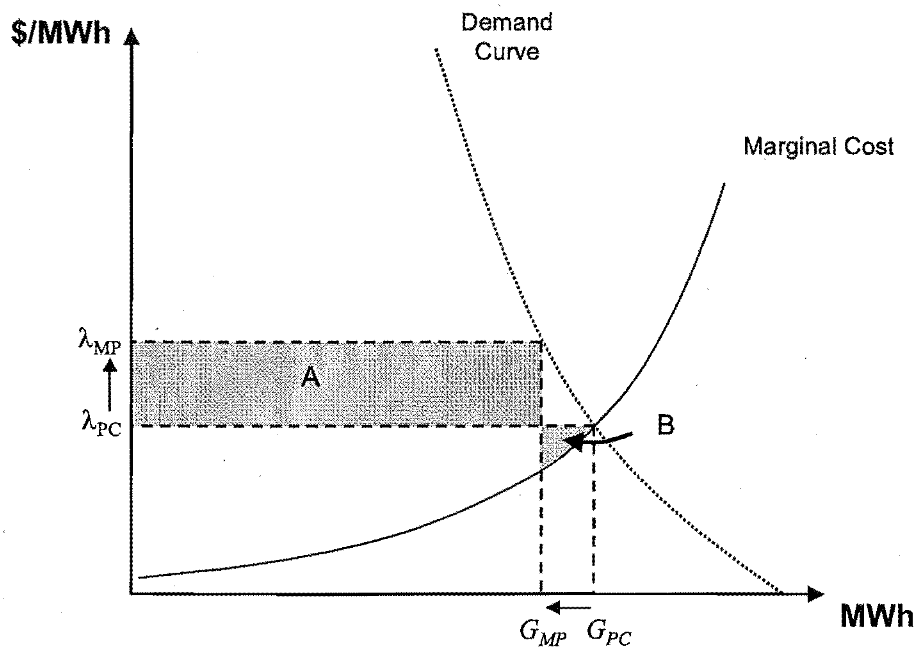


Figure 7.1 Gains from reducing output.

When the firm has sold a certain quantity of fixed-price long-term contracts, however, the increase in spot price only applies to those units sold at the spot price, i.e., the difference between total output and contract quantity. It can be seen from Figure 7.2 that this significantly reduces area A, and will result in the difference between the two areas being maximised, defining the profit maximising output and price, much closer to the competitive solution. Since the marginal revenue to the firm is described by the *change* in area A, as the firm changes output, a positive quantity of fixed-price contracts causes a leftward shift in the marginal revenue curve (Wolak (1999)). Scott (1998)'s model of a Cournot electricity market (see Section 7.5 for a full consideration of the model) showed that if a firm has sold a quantity of contracts equal to the competitive output, the competitive outcome will be observed in the spot market. Again, this is intuitive, since in the case where all output is sold forward, the marginal revenue curve is constant at the level of the spot price, for all output levels.

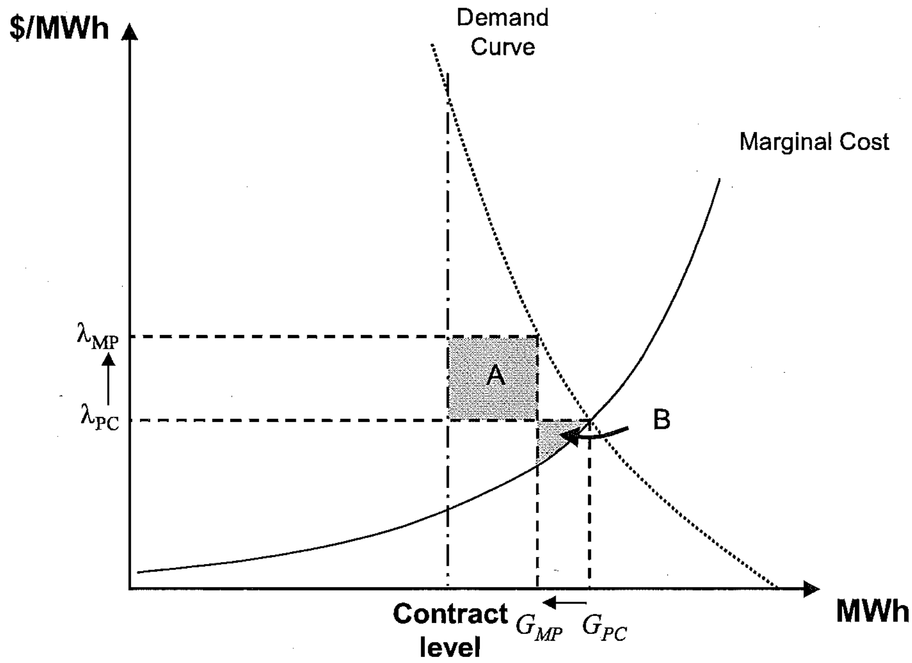


Figure 7.2 Output decision with contracts less than competitive output

Furthermore, as shown in Figure 7.3, a contract quantity large enough (greater than the competitive output) will result in prices below marginal cost. Here the firm is effectively a net buyer from the spot market, in order to meet the contract quantity, and thus has an incentive to push spot prices down. The firm is now maximising the difference between additional generation costs, and the reduced cost of effectively buying back energy from the market (Figure 7.4).

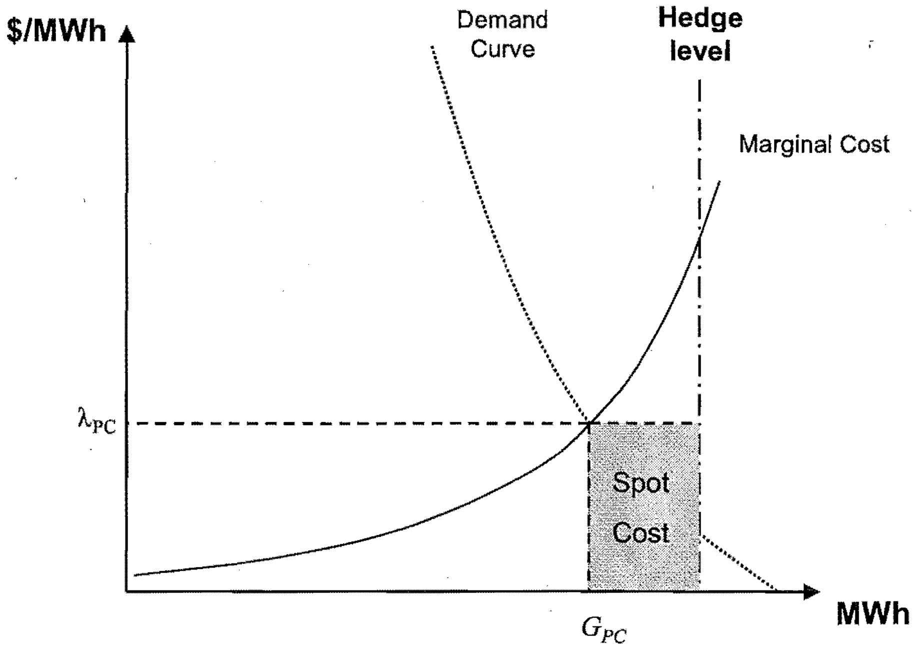


Figure 7.3 Output equal to competitive level, contracts higher than competitive output

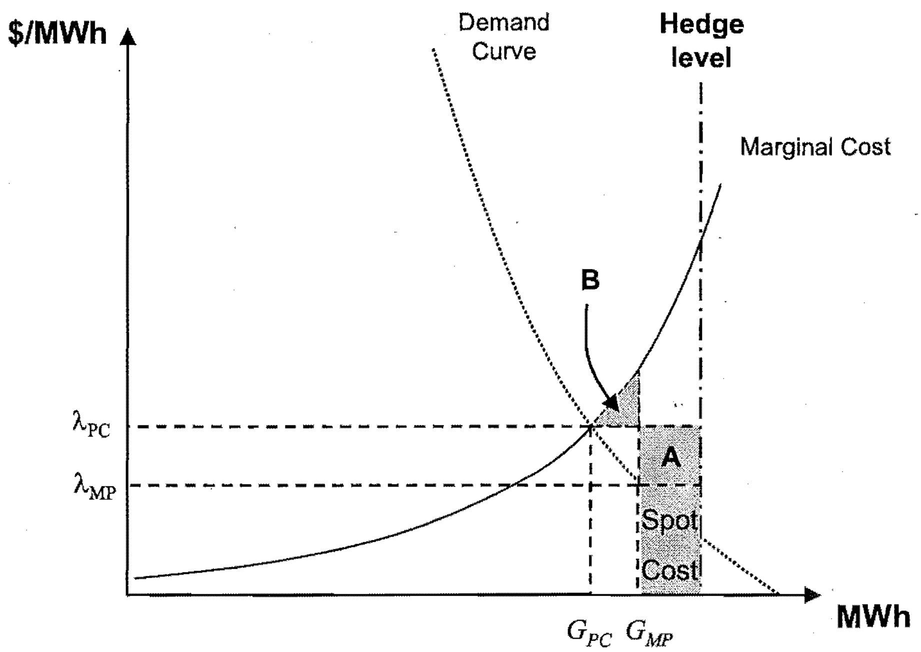


Figure 7.4 Profit maximising output with contracts greater than competitive output

Scott summarised these results into the following general relationship between contracts and output decisions by a firm with market power:

> *"The effect is to 'distort' output towards the contract quantities....in general terms, production will fall between the contract amount and the amount a perfect competitor would produce at the prevailing spot price."* Scott (1998), p 29 and 34

Such a result gives rise to the general rule that contracts reduce the incentives for firms to exercise market power, and result in solutions much closer to the competitive one[26].

7.4.1.2 Joint Spot-Contracting Optimal Policies

While Scott provided results for the effect of contracts on market power, he did not investigate the implications for optimal contracting strategies. Although he implicitly modelled risk neutral generators, the model provides insights for risk hedging. Results from the model, outlined in the next chapter, show that for low levels of contracts, overall profit variability was well below that which would be observed if the firm behaved competitively. This implies that market power itself provides a "natural" level of risk hedging for an individual firm through the elasticity of residual demand. A decrease in output is offset (in revenue calculations) by an increase in price. However, the relationship between contracts and spot price variance was complicated. This aspect of the contract-storage-risk relationship will be developed further in Chapter 8.

Scott's results possibly raise more questions about the contract-market power relationship than they answer. It would appear that for consumers of electricity, not only do forward contracts with suppliers hedge against spot price risk on the amount covered by the contract, they also predispose oligopolistic supply firms to a higher level of output and lower spot prices for the amount not covered by contract, reinforcing the incentives acting on consumers to buy contracts. On the other hand, generators face a reduction in spot profits as they increase contracting, as their output approaches the competitive

[26] An interesting note here is that these results appear to mirror the more general analysis of Brander and Lewis (1986), who investigated the effect of the debt structure of a oligopolistic firm on their performance in the spot market. The authors found that leverage acted as a pre-commitment to a particular spot market strategy

equilibrium. While it may appear that risk averse generators would sacrifice some spot profit in order to achieve profit stability through forward contracts, they could have achieved at least some degree of stability through the exercise of market power.

Hence the question arises as to why supply firms in this situation would sell any proportion of their output on hedge contracts at all. Analysis of this question requires a representation of the contract market and the incentives present for all contract market participants to hedge, not provided by Scott (1998). Two strands of literature have focused on optimal hedging strategies given oligopolistic behaviour in the spot market, and yielded two rationales for contracting in imperfect markets: to be able to strategically influence your rivals' output strategies, and to make contract profits where consumers' risk aversion can be exploited in contract prices. The majority of these analyses use the common two-stage decision framework where it is assumed that contracting decisions are made in the first stage, based on the expectation of optimal spot market outcomes given a level of contracts. This is a reasonable approximation to an electricity market, where delivery is instantaneous, and forward positions are set well in advance of the delivery period.

A two stage model of oligopolistic spot-contract equilibria was first introduced by Allaz (1992), for a general commodity market. Allaz's model was developed with the agricultural setting in mind, so the form of the contract was a futures contract, which were traded between the oligopolistic firms and futures speculators, the latter either taking physical delivery of the commodity or having an offsetting position with someone who would.

Spot demand is uncertain at the first stage, but this uncertainty is resolved by the second stage, when producers come to make their production decisions. Hence the expression for the optimal production decision, i.e., the first order condition of the second stage fixed-contract[27] profit maximisation, is found to be similar to the form of those developed by Scott (1998) (i.e., a firm's output was increasing in the level of contracts sold). This expression, for each supply firm, is a function of both the firm's own contract level, and

[27] Since contracts were set in the first stage of the model

the output decisions of its rivals. Hence the system of first order equations (whose solution represents the equilibrium at stage 2) implicitly defines optimal output as a function of the aggregate forward position of all supply firms. More specifically, rivals' output was decreasing in the number of forward contracts sold by a firm.

At stage 1, producers and speculators determine their forward positions, given their aversion to the uncertainty of the exact spot outcome in stage 2, and their knowledge of how the spot equilibrium behaves with respect to contract position. Two strategic relationships between contract and spot decisions were present. Firstly, Allaz introduced a single "conjectural variation" parameter, effectively representing the nature of quantity competition in the contract 'market'. More exactly, the parameter measured the (expected) change in a rival's contract quantity for a unit change in the firm's contract quantity. Secondly, a firm was able to indirectly affect a rival's output decision through the knowledge of the equilibrium conditions in the spot market, as discussed above.

Allaz evaluated the resulting expressions for the equilibrium contracting strategy for a range of producer risk aversion, and conjectural variation. One of the more interesting results showed that even when the producers were risk neutral, a positive level of contracts was still traded in equilibrium. Since the producers had no risk hedging motive to sell forward, these contracts highlight the strategic motive for contracting: namely that an individual firm can influence its rival's contracting and output strategy. When two firms competed *á la* Cournot in the contract market, a firm would choose a contract level that was just over half the no-contracting duopoly output level. However, when the contract market equilibrium was considered, this did not have the effect expected by the firm:

> *"For an individual producer, this policy could pay off. However, when they all do so, ignoring each other, they end up producing too much and are worse off."* Allaz (1992) p 304

The policy Allaz refers to only 'pays off' if it is assumed that one producer can act in anticipation of the others, i.e., a Stackelberg situation. It is also interesting to note that while the firms know that all producers will operate according to the same first order conditions in the spot market, (i.e., they are aware of the relationship between contracts and the resulting spot equilibrium output), Allaz assumes they are ignorant of the fact that other firms will perform the same contract optimisation in the forward market.

Allaz and Vila (1991) included a number of trading periods, where firms could sell forward contracts for a specific production period, in any of n periods prior to the production time. They found that as n tended to infinity, under the assumption of perfect foresight, that the market would ultimately become competitive. This result was highly dependent on the structure of the game – that each period, firms would simultaneously choose an adjustment to their previous aggregate contract level, and these contract quantities would be subject to a Bertrand 'auction' to determine the price. As discussed above, it can be shown that in this situation, an individual firm can profit from increasing contracting, as long as its rival does not change its strategy. If both firms performed the same optimisation, they were worse off. The interesting aspect of this model was that each firm kept pursuing the Stackelberg strategy as more trading periods were allowed. While the result that increasing n to infinity (and thus allowing more contracts to be sold) will ultimately lead to the competitive output is mathematically true under this structure, it seems unlikely that firms would not 'learn', even over a small number of periods, that its rival is performing the same optimisation.

Hughes and Kao (1997) considered the case where firms are not aware of each others' contract position, to see whether the strategic rationale for contracting still existed. Firms expect risk-averse rivals to enter into forward contracts to hedge demand uncertainty, even if these contract positions are unobservable, and so the incentive to influence their contracting strategies remained.

Gans, Price and Woods (1998), while not discussing optimal contracting per se, considered the effect of a contract market on the entry of new supply firms, by allowing a potential entrant as well as the Cournot competitors to sign pre-emptive long term contracts. While the static, single period Cournot models of Scott (1998), Allaz (1992) and other authors ignore the relationship between resulting prices and the attractiveness of entry, Gans *et al* determined the post-entry prices (and profits) in a Cournot framework, both with and without a contract market. The aims of the study were to determine if an efficient entry decision (i.e., the potential entrant had a marginal cost lower than the average marginal cost of the incumbents) would be changed by the existence of a contract market, since it was asserted (in similar fashion to the earlier work of Allaz) that contracts predispose firms to lower prices and therefore profits. A firm would find entry attractive if its expected post-entry profit covered the fixed sunk entry cost it incurred. Obviously, for extreme levels of this sunk cost, the presence (or not) of a contract market would not affect the decision of the entrant – extremely high entry costs would deter entry, and extremely low entry costs would trigger entry regardless of the effect of contracts on prices. However, it was shown that there existed an intermediate range of fixed entry costs where efficient entry would have occurred, were it not for the presence of the contract market, which caused lower prices and thus lower profits. However, the implicit assumption that the entrant did not have access to the contract market, and thus would be completely exposed to the spot market, means that the game being played was essentially asymmetric. While the incumbents had access to two markets for energy (i.e., the spot and contract market), the entrant could only compete in one. One would expect that if the entrant was able to sign contracts at a premium to the spot price, entry would become more likely in the intermediate range of prices. Also, in a more general multiperiod setting, i.e., with demand growth (not considered by these authors), we would expect to see contracts delay efficient entry, rather than prevent it, for a similar intermediate range of contracts.

The assertion that contracts lower industry profits and delay efficient entry is significantly influenced by the assumption made in the above analyses that contracts will be sold at a price equal to the expected spot price. Thus risk-neutral potential entrants will be indifferent between selling output on the spot market or selling it on the contract market.

However, the presumed risk neutrality of incumbents and entrants may be unrealistic. While it may be reasonable to believe that established firms may be risk neutral, the same may not be true of potential entrants. In the capital-intensive electricity market, the significant sunk entry costs faced by firms contemplating entry would inflate the "risk" arising from uncertain market outcomes post-entry. Additionally, a new firm would want to establish itself amongst consumers, and contracts are an ideal avenue for strategically 'locking in' customers. This would generate a higher level of confidence in its ability to survive in the early years of operation. These extra risk-aversion and strategic incentives acting on potential entrants should be addressed in any model of entry in the presence of a contract market.

It could also be argued that in many markets where forward contracts are sold to risk averse consumers of the commodity, contract prices would include a risk premium (see Chapter 5) and thus forward prices may be biased upwards (depending on the relative risk aversions of producers and consumers). In such a case, the profits made by supply firms from contract sales might outweigh the lower spot profit resulting from higher output and lower prices, providing a profit making incentive to contract, as well as the strategic and risk hedging elements. Even in the absence of consumer risk aversion, loads may be strategically motivated to pay a premium over the expected post-entry spot price, simply to encourage entry. While this may represent higher costs in the short-run, the decreased market concentration and significant additional system capacity may be sufficient reason to expect lower future prices once the new firm is established.

While this study does not intend to provide a detailed model of entry decisions, the notion of 'normal backwardation', i.e., the potential for contracts to be sold at a premium to the (expected) spot price, is of great interest.

Green (1993) and Powell (1993) applied the two stage model to the British electricity industry. The fact that contracts in this market are not standardised, transparent or 'traded' in a market with some risk-neutral participants, led these authors to allow the contract price to differ from the expected spot price. Buyers of contracts were assumed to be risk averse consumers of electricity (Regional Electricity Companies, or REC's), modelled again with mean-variance preferences. Powell's analysis was unique in that he assumed a small number of REC's competing for electricity on the spot and contract markets *á la* Cournot. In this way, REC's were aware of the influence their choice of contract strategy would have on future spot prices. If the generators acted according to Cournot conjectures in the spot market, and Bertrand conjectures in the contract market, the latter drove contract prices down to the expected spot price, and the expected spot price to marginal cost (as firms undercut each other to gain a full share of the market).

Since the firms were identical, "equal shares full hedging"[28] was the equilibrium in the contract market, and as a result the competitive spot outcome was also observed, since generation firms no longer had any incentive to game the spot price. However, when collusive behaviour by generators in the spot and/or contract market was modelled, forward premiums emerged, as the firms used their monopoly power to select the optimal level of hedging. Powell's model potentially understates the desire of demand-side consumers to purchase contracts, since he assumes that only RECs with fixed and certain demands will wish to purchase contracts. Demand side firms who exhibit price-responsive behaviour (thus providing the slope of the demand curve) were assumed not to contract. Chapter 9 will show that the inclusion of these firms in contract demand may reduce the risk premium firms are willing to pay for contracts, and increase the overall level of hedging in the market. However, Powell's general conclusions, i.e., that profit-making incentives do indeed exist for generators with market power to sell forward contract in a market of this nature, were valid. Powell also noted, with respect to entry:

> *"Non-transparent contracts imply that entrants cannot easily calculate the true present value of entry...Non-transparency may allow the generators to capture rents in the contract market, and gives the generators the ability to price-discriminate. The generators appear to have strong incentives to maintain this form of contracting."* Powell (1993), p 452

Of course, this is predicated on the fact that the generators are, in fact, able to contract profitably, and that this position is able to be supported. One would expect that the low, entry-deterring, spot prices would eventually effect the re-negotiation of contracts, possibly compromising the generator's ability to discriminate to the same degree.

A model more closely aligned with the demand side characteristics of New Zealand is found in an analysis of the British market provided by Green (1993) [29]. Green comments that the Bertrand outcome described by Powell (1993) may be unrealistic, as contract negotiations may be drawn out over a period of time, as opposed to the fierce

[28] Contracts equal to expected generation

[29] I.e, Green assumes the REC's have no knowledge of the effect of their contract decisions on spot price.

undercutting that produced the competitive outcome. For this reason, Green chooses to model contract competition with Cournot conjectures on the part of the generators. With risk neutrality (and thus no forward-spot price spread), he shows that:

> *"The effect of even the most limited competition in the contract market will increase output in the spot market by up to 20%, and reduce the gap between price and marginal cost by up to 40%...The presence of an uncompetitive contract market has produced a substantial increase in welfare, and the gains from the more competitive market that actually exists will be even greater."* Green (1993), p 6

When risk aversion is introduced by way of a mean-variance optimisation by both sides of the market, the net effect is unclear. While risk aversion on the part of the generators leads them to desire to sell more contracts in equilibrium, risk aversion by REC's drives forward prices above the spot price, and could either raise or lower the number of contracts sold in a Cournot equilibrium, depending on the degree of risk aversion:

> *"If the REC's are buying many contracts, then the generators would do well to raise the price, while if the number sold is low, a price cut to increase quantity would be more profitable."* Green (1993) p 7

An interesting extension provided by Green is to a multi-period setting, where REC's naively believe that the spot price in the next period will be equal to that in the current period. This provides an incentive for generators to sacrifice current period profits by keeping the spot price high, in anticipation of greater contract profits in the next period. Green introduces contracts to an earlier supply function analysis of the British market (Green and Newbery (1992), see Chapter 4 for a discussion of this and other supply function models). In Green (1996), he modelled both spot and contract market equilibria with supply functions, and showed that, in the limit, the contract supply functions could, at their extremes, represent Cournot or Bertrand beliefs about competition in the market, reflecting results obtained earlier about contract market equilibria under these conjectures. Significantly, Green showed that the strategic incentive for contracting, suggested by Allaz, disappeared when the less-competitive supply functions were introduced to the spot market. As Allaz had noted in his paper, the strategic motive for contracting seemed to be particular to Cournot assumptions in the spot market.

In his 1999 paper, where optimal contract positions were represented by fixed quantities, similar results were obtained again, although he was able to extend them to the case in which firms hold asymmetric beliefs about each others conjectures relating to contract market competition. While Green's aforementioned supply function analyses assumed rational expectations about the forward price, in Green (1999) the process by which risk-neutral generators could exploit consumers' risk aversion to obtain high contract profits (forward price greater than expected spot price) was also described.

Bunn, Larsen and Dyner (1997) extended Green's assertion that generators may act to influence future contract prices through controlling the mean spot price, to include the effects that spot price volatility may have on contract prices. In markets where contract prices include a risk premium, higher contract prices may result from higher volatility. Bunn *et al* used a system dynamics approach to modelling the interactions between spot and contract markets, and incorporated aspects of the British energy industry into the model. Dominant generators influenced spot volatility, deliberately withholding capacity from the market for short intervals within some time interval, thus providing a price "spike" above the short-run spot-profit maximising solution. In a market where the standard deviation of the spot price is normally around 50% of the mean price, the increase in average price from this strategy provides extra spot profits while remaining relatively unnoticeable. Generators also profited from this strategy in the contract market, in two potential ways: risk averse consumers were willing to purchase a greater quantity of hedge contracts, and possibly at a greater contract price. Not only does this lead to higher contract profits for the dominant generator, the authors argue it may result in a barrier to entry as potential independent power producers are discouraged through large fluctuations in the market.

Results from running the simulation model for various destabilisation strategies showed gains in revenue of up to 11% were possible for strategies inducing a 5% increase in price variance, and that the relationship between these two factors was relatively linear. The exact fashion in which dominant generators create volatility is likely to be more sophisticated than the somewhat systematic manner in which Bunn *et al* model, but the general insights remain the same, in particular, the fact that destabilisation behaviour may be relatively undetectable:

> *"As volatility is already high at such [peak demand] times, the potential for even a small amount of tactical behaviour of the sort indicated in this article will be very hard to notice, but potentially quite profitable."* Bunn, Larsen and Dyner (1997), p 286

7.4.2 Empirical Analyses

Helm and Powell (1992) discussed both directions of the market power-contracts relationship for the British supply industry, which is dominated by two major coal-fired generators, National Power and PowerGen. Using data from a number of years of operation of a deregulated market, they performed co-integration analysis on the pool purchase price (PPP, the equilibrium price for electricity before capacity and loss-of-load charges are added) and the demand for electricity. In a statistical regression, they used a dummy variable to represent a significant event in the market, namely the expiration of a significant quantity of the generators' vesting contracts. The dummy variable highlighted a significant (in the statistical sense) change in the relationship between pool price and demand at the time these contracts expired. The significance of these variables indicated a strong effect of the level of contracting on the pool price, in particular, it supports general theoretical conclusions (presented below) that high levels of CfD's appeared to control the degree of market power the generators could exercise, leading to lower spot prices. When the existing contracts expired, prices unambiguously rose. At that point, an incentive for electricity consumers to buy contracts, other than risk-avoidance, is illustrated:

> *"In a sense, if the regional electricity companies were to buy further CfD's, what they would be doing , apart from buying a hedging instrument, would be buying a device that controlled the market power of the generators or, rather, bribing the generators not to abuse the dominant position" Helm and Powell (1992)*, p 102

Thus it would appear that those firms that can exercise market power would not want to prevent themselves from doing so by selling a large quantity of contracts. However, as the authors note, incentives do exist for dominant firms to sell contracts. As there is little information required by market authorities on the details of contracts held by firms, the presence of a contract market "seriously devalues the transparency of the spot market" for two reasons. Firstly, it allows generators to achieve a level of price-discrimination between contract buyers. Secondly, the pool price no longer reflects the true fundamentals of the industry:

> *"Our argument here is that the pool price has risen due to a set of contracts expiring: in other words, due to nothing fundamental about the industry in terms of costs or demands or future demand or supply problems"* Helm and Powell (1992), p 102

Lowrey (1997) extended this analysis to include the second expiry of contracts, and also other significant market events, such as threats and announcements by regulators. Lowrey found that the second expiry was not as significant as the first (in terms of a change in the price level) and reasoned that this was because of the effect of increased entry by independent power producers and nuclear generation (effectively a competitive fringe) in the interim. He went on to relate other significant price level changes to the credibility of regulatory threats.

Rather than attempt to model gaming using traditional theoretical analyses, Wolak (1999) provides an empirical model of "best response bidding strategy" in each half hour by firms in the Australian National Electricity Market (NEM). Firms submit an offer curve (quantity supplied by a generation unit, and the price for that unit), and Wolak provides a definition of a Nash equilibrium in the market, where all firms maximise profits given the strategies of other players, their own generation portfolio and level of contracts, and validates the model with actual outcomes observed in the market. With this model, Wolak set out to explain why market prices in the NEM fell significantly from $AU30 per MWh just prior to re-structuring (and the introduction of contracts) to $AU15 per MWh three years later, a level roughly equal to the marginal cost of large, baseloaded thermal plants. The results supported the above theoretical analyses that report high levels of contract cover will cause a firm to bid at close to, and possibly below marginal cost, in order to dispatch as much of their capacity as possible. Low pool prices are desirable so long as generators are effective net buyers from the market (through the CfD difference payments). Wolak goes on to speculate as to why generators sold such a high level of contracts, and the reasons postulated are system specific, and may be influenced by political considerations. However, he does comment on the long-term effects of high levels of contracts, especially with respect to contract renegotiation:

> *"This low-risk contracting and bidding strategy can have dire longer-term consequences if very low market prices are necessary for the generators to sell all of its contract quantity. These low prices cause purchasers of contracts to form expectations of very low futures prices, which makes it difficult for the generator to sell future hedge contracts at prices above the generator's marginal cost."* Wolak (1999), p 38

In summary, these three empirical works appear to support the general conclusions of the theoretical papers presented in Section 7.4.1, although they stop short of modelling the effect of spot outcomes on the re-negotiation of contract prices.

7.5 Storage, Market Power and Contracts

Allaz (1991) extended the earlier monopoly work of Brianza, Phlips and Richard (1987) to the case of oligopoly, in particular, where participants competed according to Cournot conjectures. Of note here was the examination of the relative roles of forwards and futures on output decisions. Under a two-stage framework where production occurs in both periods, but spot transactions only in the latter, storage affected second stage spot market decisions in a similar way to that of forward contracts outlined by Allaz (1992), where the value of a strategic mechanism is chosen in the first period to alter the point at which marginal cost and revenue equivalence was found in the second. Intuitively, we expect that production for the sale period will be split evenly over the two periods (ignoring discounting), so as to minimise total production costs (which were quadratic). However, production (and thus storage) was higher in the first period than the second, diminishing production costs in the second, again leading to a higher output on the spot market:

> *"Forward sales made at time 1 affect the marginal revenue curve for spot sales at time 2. Inventories carried over from time 1 affect the marginal cost curve at time 2. In both cases, the producer is less price sensitive at*

> *time 2 on the spot market and, hence, tends to increase his output."* Allaz (1991), p 265

However, when storage and contracts were simultaneously available, the forward market was the avenue used for this strategic manipulation. This was due to the quadratic form of the marginal production cost function, as any deviation from producing the same quantity in each period (all other things held equal) resulted in a higher total cost over both periods. Hence, even in the absence of storage costs *per se*, storage in this manner "costs", while contracts are relatively costless, under Allaz's assumption that the futures price was equal to the expected spot price. One could also question the relevance of the two-stage model to a multiperiod setting. While it seems realistic that contracts for a given period will be determined well in advance, production in the first period is not solely for storage, at least in an electricity market. In other settings, where production and sales may be offset, this model may represent the incentives acting on the firm at each time period. However, for a hydro firm, it must trade off water for release in a period against water stored for a future period. While the hydro firm may be less 'price sensitive' in time period 2 if it stores more, it will be more price sensitive in time period 1 as a result.

Scott (1998) modelled a duopolistic market in the context of a hydro reservoir management model. As described in Chapter 6, Scott's storage policy was defined recursively by dual dynamic programming. An end of period demand curve for storage (or DCS, pairing water values with desired storage position at that value) was added to a demand curve for release (DCR). This process equates to trading off the marginal value of storing water against the marginal benefit from releasing it now, analogous to economic arbitrage. Unlike Read (1989), where the DCR was effectively a stepped supply curve representing the prices at which thermal stations would be baseloaded in a centrally planned system, Scott's DCR needed to model the strategic response by competitors to a change in quantity released by the hydro firm, and hence the DCR was downward sloping over sections where the rest-of-industry marginal cost curve was flat. The spot market was modelled as a Cournot duopoly with a fixed volume of long term contracts, similar to the second stage of the models of Allaz (1992), Green (1993) and Powell (1993). This "single period model" was run for different values of constant marginal costs, representing different water values of the hydro firm, and yielded the market outcomes for each, in this way defining the DCR. Adjusting for the uncertainty in (uncorrelated) inflows, the beginning-of-period DCS was achieved. Through backwards recursion, a water value surface was obtained which described the optimised water values for each storage level, and period, in the year. This was then used in a simulation of a number of different actual inflow sequences.

Although the model took account of the incentives from a fixed level of forward contracts there was no market for the trading of the contracts themselves. However the simulation model was run for a variety of contract levels to observe the impact on system outputs such as price, generation, storage and profit. The results show that the level of contracting has a significant impact on these outputs:

> *"Our experiments...lead us to conclude that the single biggest effect on price distortion is the level of contracting, with generation rising and energy spot price falling as contracts increases."* Scott (1998), p 134

Results for generation, storage and energy spot price from Scott's model are summarised in Appendix B, while the important results for profit risk are examined in depth in the next chapter.

7.6 Conclusions

The purpose of recent chapters has been to establish what insights the existing literature has for the decision problem outlined in Chapter 3. Recalling the important aspects of the problem, a hydro manager may use water storage, long-term forward contracts (in a relatively illiquid market) and market power to manage risk and return in a mixed hydro-thermal electricity industry. The literature summary first considered these three tools separately, and then established results (if any) for their interaction, and its effect on risk management. The key relevant insights, for the hydro manager's problem as described in this thesis, are briefly summarised in Table 7.1 below.

	Interactions of		
Effect on:	**Storage and Contracts**	**Storage and Market Power**	**Market Power and Contracts**
Return	(A) Few relevant insights	(C) Traditional inter-temporal water trading incentives even stronger for firms with market power; higher than competitive returns possible under certain assumptions (Dalziel (1987)) These traditional incentives may act in the opposite direction in some mixed hydro-thermal systems (Bushnell (2000))	(E) General result that increasing a firms contract level shifts output and price towards competitive solution. (Scott (1998), et al) However, contracts may be used to strategically influence competitor behaviour (Allaz (1992)) Spot market behaviour may also be used to increase price expectations and thus contract prices (Green (1993))
Risk	(B) Few relevant insights	(D) Reinforcing of traditional water trading policies act to further stabilise prices under uncertain inflows (Newbery (1984), Dalziel (1987)) Again, in some mixed-hydro-thermal systems, prices may be deliberately spiked during high demand periods (Bushnell (2000))	(F) Use of spot market destabilisation, to increase consumer perception of risk and thus risk premia, postulated (Bunn, Larsen and Dyner (1997))

Table 7.1

Analyses of the interactions between all three are scarce. Brianza, Phlips and Richard (1987) suggested that contracts and storage are substitutable tools for inter-temporally shifting marginal revenue curves, but that contracts are preferable because they are less "costly" than storage, under the assumption of positive storage costs. Given the general results in boxes C and E, one could postulate that contracts, however, might negate some of the stabilisation effects of market power suggested by Dalziel (1987) and Newbery (1984). Since forward contracts shift output towards the competitive solution, thus increasing release, one would deduce lower levels of storage and a higher susceptibility to inflow variations would be observed. This contract could be viewed as 'riskier', as the firms' uncertainty about being able to meet contractual commitment grows with the level of contracts. Chapter 8 will attempt to verify this, and other aspects relevant to boxes A and B, with an empirical analysis of the gaming hydro model developed by Scott (1998)

Many of the above analyses derive expressions and results for the effect contracts have on spot market outcomes, and deduce that, under the assumption that contract prices are equal to the expected spot price, contracts will only lead to lower expected profits for firms with market power. However, examples abound where dominant electricity firms have chosen to sell a significant proportion of their output on contract Wolak (1999). While in certain cases this is as a result of explicit or implicit regulation, we believe that in many cases firms are able to earn additional profits through selling contracts at prices greater than the expected spot price, which they in turn have a degree of control over. Few authors (possibly only Green (1993) and Powell (1993)) provide comprehensive models of these dynamics in electricity markets, as Scott acknowledged:

> *"A major dynamic factor we have not captured in our model is the way in which contracts are re-negotiated over time. It is reasonable to expect that prices today will influence both the price consumers will pay for contracts (in the medium-term) and the overall demand for electricity (in the longer term)"* Scott (1998), p 136

Filling this void in Scott's work represents an important aim for this thesis. While Green (1993) and Powell (1993) attempted this, we believe some of the critical assumptions employed were unrealistic for the market we intend to study, namely:

- The way in which consumers form naïve expectations of future spot price, i.e., next year's spot price will be equal to this year's spot price.

- The implicit belief that while contract prices are driven by risk aversion and thus the variance of spot prices, consumer's measurement of spot variance cannot be influenced by dominant firms through spot market strategies

The results and insights from Bunn, Larsen and Dyner (1997)'s system dynamics model provide an important motivation for further analysis of the profits available from risk management of a different kind: increasing the risk for other market participants, consumers in particular. We intend to provide further support for these results, with a more analytical representation of the spot and contract market.

Having concluded our review of the literature, we now turn our attention to an analysis of the risk faced by generation firms in a hydro-dominated market, and then to developing our model of joint spot-contract equilibria.

8

HYDRO RISK

8.1 Introduction

The previous chapter summarised the established principles in the literature on the management of "market risk", defined as the variance of profit, by a dominant hydro firm that has the opportunity to sell long-term contracts. It highlighted the need for further investigation into risk management in this context, as the vast majority of existing studies have not simultaneously considered the three mechanisms illustrated by the risk management triangle described in Chapter 3.

In particular, Chapter 7 posed questions relating to the effect of hydrological uncertainty on a hydro firm's profits. At first glance, it would seem reasonable to expect that long-term contracts and storage might perform well together in managing the risk associated with variable inflows, given that they are each individually known as instruments of stabilisation. Additionally, the ability to exercise control over market outcomes through a dominant market position could also be seen as stabilising.

However, many authors have noted that high levels of contracts may over-commit a firm whose ability to supply is uncertain. Under the arrangements of a contract for differences, if a firm cannot (or chooses not to) meet all or any of its contract obligations, it effectively purchases the shortfall off the spot market, at the prevailing spot price. As an extreme example, a high level of contracts, combined with a low generation capacity and high spot prices would be very costly for the firm. In general, the risk the firm is exposed to is a combination of three critical factors: the contract level, the uncertainty as to its ability to produce those contract commitments, and its uncertainty about, and/or influence over, the spot price.

Any electricity generation firm will face uncertainty in its ability to supply, due to the possibility of plant failure and/or transmission congestion. Firms may also face significant uncertainty in their generation input. While a thermal firm may be reasonably certain about its supply of fossil fuels, in most countries with hydro capacity, a hydro firm will face a moderate degree of inflow uncertainty. The firm's ability to store inflows will determine how much it can "smooth out" inflow variability (whether these are unpredictable fluctuations or predictable seasonal variations), and thus determine the degree of certainty the firm has as to its ability to generate in a given period in the future.

The situation is further complicated in markets that are dominated by hydro generation. Years of high spot prices are very likely to be correlated with drought years, when all hydro firms face the same meteorological conditions[30] and are thus restricted in their ability to generate. A firm that has a low level of contracts may find this correlation has a stabilising effect on profit: while its output in drought years will be low, the price it receives for its spot output will be high. However, this correlation may present significant risks for a firm with a high level of contracts. At the very time that the firm is least likely to have the capacity to meet its contract obligations, it will also face high spot prices and thus significant costs in making its contract difference payments.

While non-hydro firms may not face the same degree of supply uncertainty, and its correlation with the spot price, being a participant in a hydro-dominated system presents an extra degree of uncertainty in the spot price driven by hydrological conditions, in addition to load uncertainty.

---

[30] As is likely to be the case in geographically small markets, such as New Zealand

While the traditional wisdom reports that selling forward contracts provides greater certainty and thus reduces risk, it is clear that the situation is more complicated for a firm that faces supply uncertainty. This chapter addresses these issues. More specifically, for a firm that maximises profits and competes *á la* Cournot with a single rival, we investigate:

- The extent to which the variability in hydrological inflows is reflected in profit variability for the hydro firm,
- The relationship between contract level and the variance of profit, and
- Whether, for any level of contracts, profit variability would be concerning to a risk averse hydro manager.

We will address these issues in turn, in Sections 8.4 and 8.5. In order to do this, the hydro gaming model developed by Scott and Read (1996) is used to evaluate the risk position of a hydro firm via simulation. While the important details of the model are discussed in Section 8.2, two key features of this model are:

(a) The contract level is fixed, and

(b) The only stochastic element of the model is inflows.

Thus, for a given contract level, the model provides an optimisation of both reservoir management and spot market behaviour under uncertain inflows. The contract position is not optimised. Additionally, the firm is assumed to manage its reservoir and spot position in a profit maximising or risk-neutral manner. This model suits the purpose for which we intend it here: to evaluate the extent to which a firm, acting in a profit maximising manner, is exposed to profit risk when the underlying uncertainty is the quantity of inflows it will receive in a given period. The existence of risk will help us determine whether significant gains could be made from managing spot, storage and contract positions to account for risk, should the hydro manager be risk averse. If no significant degree of risk exists, for any contract level, we can assume that, for the particular system described in Section 8.2, inflow variability would not be of concern to a hydro manager, even should he or she be risk averse. In particular, the remaining tools of storage and market power would then be sufficient to manage this risk.

8.2 Market Setting

The model used in the following analysis closely resembles the New Zealand electricity industry prior to the New Zealand government's "split" of its large, state owned, generator, into three separate generation companies. This is convenient for this study, since, until the breakup, the market was essentially a duopoly with a competitive fringe. However, the hypothetical system modelled here consists solely of two generation firms[31].

Firm One has one hydro station with release capacity of 1500 MW and a single reservoir with a capacity of 3000 GWh. Reservoir inflows average 250 GWh per week[32], but the mean and standard deviation of inflows, for a given week, varies over the year. On average, the standard deviation is approximately 60% of the mean (Figure 8.1).

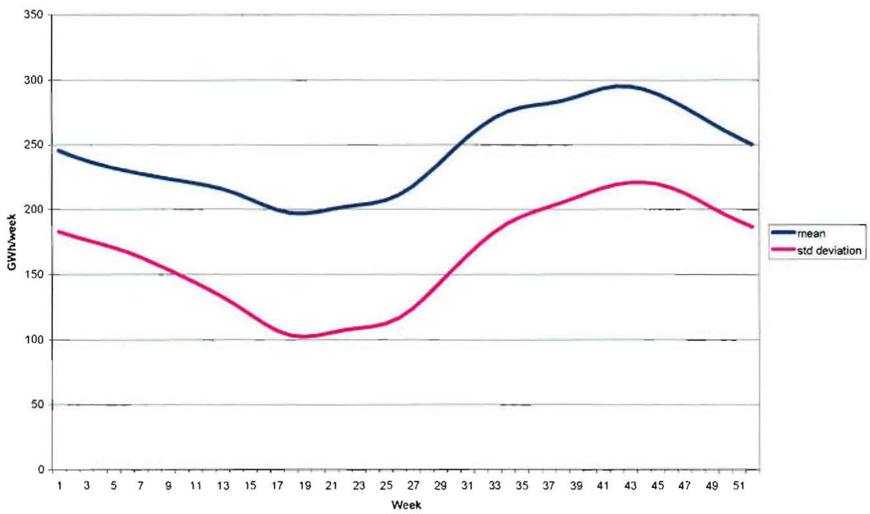


Figure 8.1 Distribution of inflows over 52 weeks

[31] Scott's PhD thesis included a competitive fringe, but this gave rise to unstable equilibria which made direct interpretation of the results, especially with regard to variance in the market, difficult. Hence we have ignored the fringe aspect of his analysis.

[32] Hence the firm can store approximately 12 weeks' average inflows

The hydro firm submits a single generation quantity to the market, with the marginal water value being its marginal cost. The water value in a particular period, as described in Chapter 6 and 7, is the result of optimally managing the reservoir given expectations of inflows, gaming and demand. The water value function will be illustrated, and its derivation discussed in more depth, in Section 8.3.

Firm Two has four thermal stations at constant marginal costs of \$10, \$30, \$70 and \$90 per megawatt-hour (MWh), each with a capacity of 750 MW. The thermal firm's marginal cost curve is illustrated in Figure 8.2.

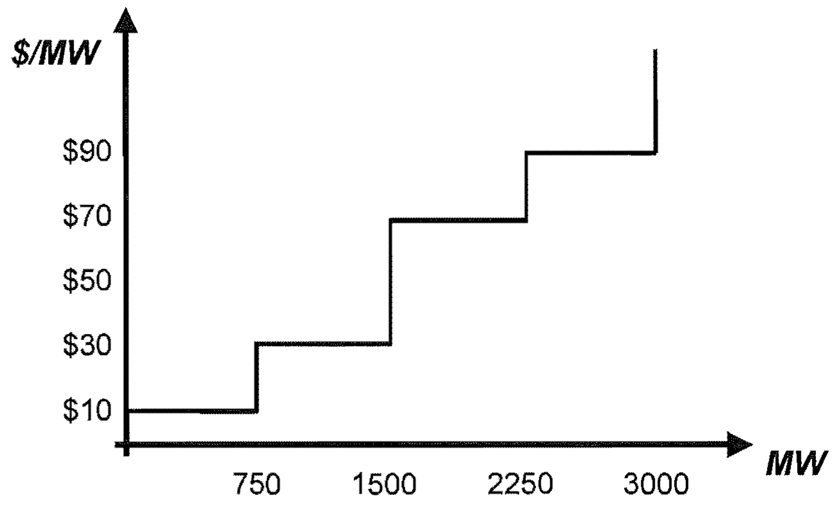


Figure 8.2 Thermal supply curve

The two firms face a deterministic aggregate load, described by the linear inverse demand curve:

$$p(G) = p_0 + \rho(G - G_0) \tag{8.1}$$

where G is the total generation from Firm 1 and Firm 2, ρ is the slope of the demand curve, and p_0, G_0 are reference points on the function. Initially, we will set $\rho = -0.005$[33], and the reference generation (G_0) will range from 3000 MW in summer to 4000 MW in

[33] Equivalent to an elasticity of 0.25 at the competitive solution

winter, at a reference price (p_0) of \$18/MWh[34]. A competitive dispatch would then yield an average spot price of \$43/MWh in summer and \$68/MWh in winter.

Neither firm was assumed to hold backup contracts, or one-way options with other firms. In the calculation of profit, contracts are valued at the mean spot price over the year. In other words, we assume that the forward price is an unbiased estimator of the expected spot price. This assumption is made for two reasons. First, it provides a reasonable lower-bound on expected contract revenue for the firm, given that it is likely that dominant firms will be able to extract risk premiums from risk averse contract customers (see Chapter 9). Second, models developed in later chapters assume the market is in a state of long-run equilibrium, in that the consumers' best estimate of the expected spot price is based on the mean spot price from a given year. In this sense, the assumption employed here asserts that consumers "got it right", and that the price paid for the contract was in fact the average spot price over the year.

8.3 Model Description

The model that will be used to examine profit variability for a hydro manager contains both an optimisation and a simulation. In each period, the firms compete with Cournot conjectures, i.e., each firm calculates its profit maximising response to the other firms' output, assuming they will hold their output fixed. This "game" equilibrates both firms simultaneously cannot improve their position, given the other firm's position is fixed.

It is also worth noting that there is a slight asymmetry in the manner in which the game is played out. While at first glance, it may appear that each period, a "one-shot" Cournot game is played out (i.e., without considering the impact on future equilibria), the hydro firm does consider the future actions of its thermal rival(s) in the way in which the water value is derived. As will be discussed in Section 8.3.2, the derivation of the water value, which in turn determines the hydro firm's profit maximising output, is the result of a tradeoff between releasing water now, and releasing water at some point in the future.

[34] Giving maximum prices of \$110/MWh in the summer and \$134/MWh in winter, and, on average, competitive demand of approximately 500GWh/week

Naturally, the value of releasing in the future is influenced by the thermal supply curve. The thermal player(s), however, make no such consideration of future prices. One could imagine a situation where the thermal player also considers the effect of its output decision on the hydro player's storage levels, and thus its behaviour in the future. However, in Scott's model, this conjecture is unique to the hydro player.

The Cournot-Nash equilibrium generation quantities for the two firms, for a given supply curve summarising the marginal cost structure of the thermal firm, a single water value for the hydro firm, and a fixed level of contracts for each firm, are found as follows. Profit, for each firm, *i* & *j*, is defined as:

$$\Pi_i = p(G)(g_i - k_i) + k_i\bar{p} - C(g_i) \tag{8.2}$$

for firm *i*, and similarly for firm *j*. Here $p(G)$ is equation (8.1), $\bar{p}$ is the average spot price for the year, and k_i and $C(g_i)$ describe the contract level, and total generation cost for firm *i*, respectively (the equivalent notation applies to firm *j*) . The cost function describes the area under the marginal cost curve: for the thermal firm, it is the sum of the areas under each segment of the stepped supply curve, while, for the hydro firm, it is simply the water value multiplied by the generation quantity.

Each firm maximises (8.2) assuming that contracts and, under Cournot conjectures, its rival's generation, is fixed. The solution to the set of simultaneous first order conditions (one for each firm, with respect to generation) describes the Nash-Cournot spot market equilibrium. Leaving aside, for the moment, the derivation of the marginal water value, the solution issues regarding the stepped thermal marginal cost curve, and assuming marginal cost is constant for both firms, the optimal equilibrium generation level for each firm can be expressed as:

$$\hat{g}_i = k_i + \frac{\pi_j - 2\pi_i - p_0 - \rho[K] + \rho G_0}{3\rho} \tag{8.3}$$

where

k_i, k_j are each firm's contract quantity, in MW

π_i, π_j are each firm's marginal generation cost.

Equation (8.3) is very similar to the Cournot equations developed by Allaz (1992), Green (1993), Powell (1993) and others, and reflects the fact that optimal generation quantities exhibit an increasing relationship with the level of contracts. Substituting (8.3) in the inverse demand relationship (8.1) above will yield the equilibrium spot price in the market:

$$\hat{p} = \frac{p_0 + \pi_i + \pi_j + \rho[K] - \rho G_0}{3} \tag{8.4}$$

It is clear that while these equations are well behaved for marginal costs that are constant across the entire range of positive generation levels, the reality of the thermal marginal cost curve, illustrated in Figure 8.2, and the uncertainty of the marginal water value, means that the derivation of equilibrium output levels and prices is much more complex. Both firms' output will vary in response to the uncertain marginal water value, which varies with storage and inflows, while the thermal marginal cost will display significant 'jumps' as generation varies outside the range of output implied by a given cost value. These aspects of the firms' costs imply two significant solution issues which were incorporated into Scott's model.

8.3.1 Thermal Marginal Cost Curve

Scott provided an algorithm that ensured that the equilibrium output level of the thermal firm was consistent with the (constant) marginal cost assumed in expression (8.3). This was achieved by partitioning the marginal cost curve into regions of generation where the marginal cost was constant (the "flats" in Figure 8.2) and regions of marginal cost where generation was constant (the "steps" in Figure 8.2), as shown in Table 8.1.

Thermal Generation	Marginal Cost
$g = 0$	$0 < \pi < \$10$
$0 < g < 750$	$\pi = \$10$
$g = 750$	$\$10 < \pi < \$30$

$750 < g < 1500$	$\pi = \$30$
$g = 1500$	$\$30 < \pi < \$70$
$1500 < g < 2250$	$\pi = \$70$
$g = 2250$	$\$70 < \pi < \$90$
$2250 < g < 3000$	$\pi = \$90$
$g = 3000$	$\$90 < \pi < \$1000$[35]

Table 8.1 Partitioning of Thermal Supply Curve

Expression (8.3) was evaluated, for a given contract level and water value, for each of the marginal cost values given in the right-hand column of Table 8.1, and the resulting optimal generation level compared with the generation bounds described in the left-hand column. Scott showed that for reasonable cost and demand functions, only one of the rows in Table 8.1 would provide what he termed an "admissible" solution, i.e., there was a unique optimal generation level which fell within the bounds implied by the marginal cost value assumed in its derivation.

8.3.2 Derivation of Marginal Water Value

Second, the model requires a way of describing the value of water in a given period. As discussed in Chapter 6, the marginal value of water should reflect not only a consideration of its use in the current period, but also its potential use in the future if the unit of water was stored.

While many traditional reservoir management policies are derived using primal dynamic programming, Scott made use of the feature that optimal reservoir policies trade water between periods (by storing inflows) so that the marginal value of releasing water is set equal to the expected value of storing water in each period[36], and used instead stochastic dual dynamic programming (SDDP). While the exact procedure for deriving marginal water value curves is described more fully in Scott (1998), it can be briefly described as follows. Solving the single period model (i.e., the solution to the system of equations

[35] This represents the upper bound on generation for the firm.

[36] Ignoring storage bounds. These were dealt with by assuming that the water value was equal to the shortage cost when the reservoir was empty, and zero when the reservoir was full (since additional water must be spilled and thus has no value)

implied by (8.3)) for a range of water values, describes what Scott termed a "Demand Curve for Release" (DCR). This is not to be confused with the demand curve for electricity (modelled by the function described in equation 8.1), but instead describes the release policy as a function of the water value observed in a given period[37][38]. Scott showed that the DCR is a decreasing function of the water value. A "Demand Curve for Storage" (DCS) describes the amount of water we would wish to hold in storage, for a range of marginal water values. Again, this is a decreasing function of the water value: high levels of water in storage imply that additional water is not as valuable to us when we have low levels of storage.

Scott's SDDP is primarily based on the recognition that release policies can be described by simply adding these demand curves, i.e., that the value, or price, of water, in any period, can be defined by simply adding the two curves which describe the demand for its use. At the beginning of a period, the total demand for water, at a given marginal water value, is the addition of the quantity required for release, and the quantity required for storage at the end of the period. Figure 8.3 illustrates this addition of demand curves. *t3* is the DCR for the current period which implies, for some water value ψ, a profit maximising release of *r1*. *t2* is the DCS at the end of the current period, which for the same water value ψ requires that *s1* be held in storage. The total demand for water is represented by line *t1*, which is the addition of the release and storage requirements at each water value (the quantity of water *s3* for water value ψ). Since the optimal output of the hydro firm is strongly influenced by the reaction of its competitor, the shape of the thermal supply curve is evident in the DCR, and less so in the DCS.

However, the demand for water within a period can also be met from another source, i.e., inflows. Hence the beginning-of-period DCS is found by adding the DCR in the period and the DCS at the beginning of the following period, as above, and subtracting inflows. The problem of uncertain inflows is dealt with by discretising the inflow distribution in

---

[37] Naturally, we think of a demand curve for a good as describing price as a function of quantity. Scott's method of creating the DCR reverses this intuition, by describing the quantity released as a function of the price of water.

[38] As pointed out by Scott, in a centrally coordinated system, the DCR is simply a stepwise function corresponding to the thermal marginal cost curve. However, in a Cournot setting, the curve is more complicated, since the thermal firm's output will vary across its marginal cost curve.

each period into 5 possible inflow "scenarios", each with an associated probability. The 5 resulting DCS's are then averaged to give an expected beginning-of-period DCS.

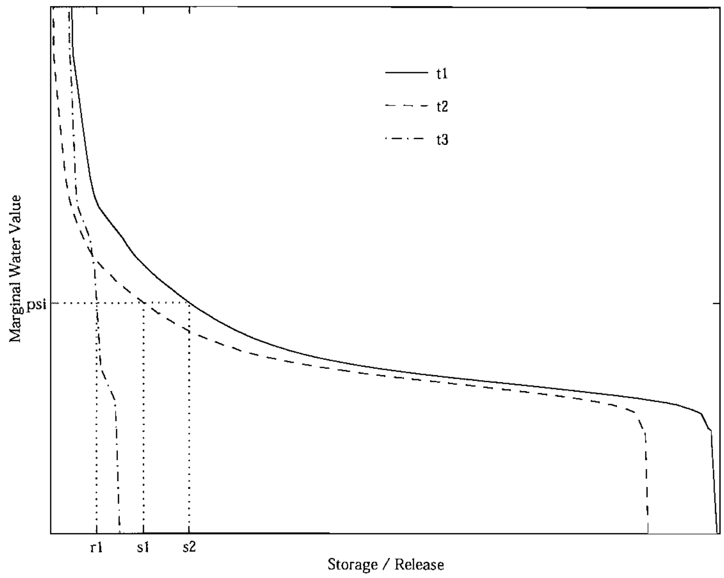


Figure 8.3 Addition of DCR to DCS, to obtain total demand for water in a period

Starting with the end-of-horizon DCS[39], the model can be solved by backwards recursion. For each period a DCS is derived that is a function describing the value of water for a given storage level, in that week. Over a full year, a "water value surface" (WVS) is created that provides the important guidelines for operating the reservoir in an optimal fashion. Most importantly, the WVS describes the marginal value of water, for a given week of the year and storage level (Figure 8.4), which can then be used to determine the hydro firms profit maximising output level. The behaviour of the WVS over the year is relatively similar. When storage levels are low, marginal water values are high and very sensitive to changes in storage. Across an intermediate section of storage levels, the water value function is very flat, indicating that moderate changes in storage will not impact the operation of the generation plant to any great extent. At high levels of storage, the value of water drops towards zero.

---

[39] Which could be derived in a number of ways. In equilibrium, the end-of-year DCS should be the same as the beginning-of-year DCS. After stepping back a full year, Scott compared these two curves, and ran the model (backwards) for a further year if they were not similar (given some tolerance). He found that equilibrium between the end-of-horizon DCSs was usually found after 3 or 4 iterations of this process.

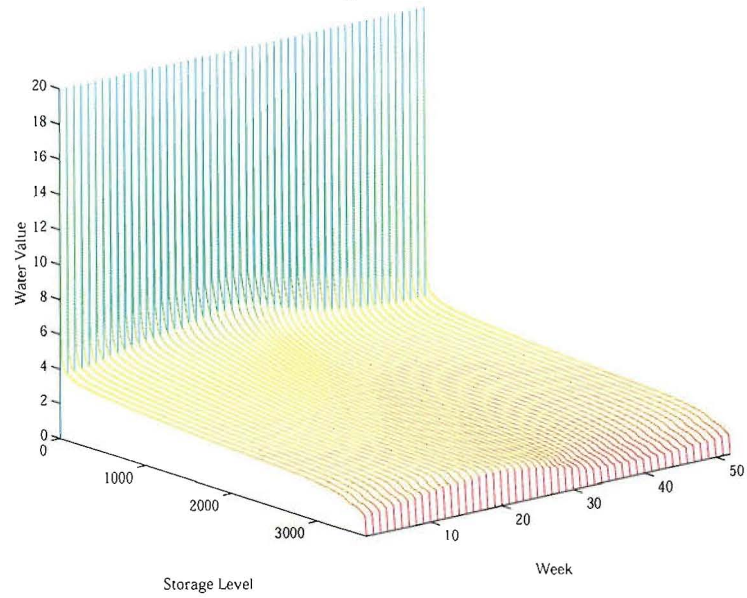


Figure 8.4 A typical water value surface

8.3.3 Simulation

The derivation of the WVS, described above, allows us to simulate the weekly behaviour of the hydro and thermal firm in the spot market over 40 years, given annual inflow sequences[40] sampled from a historical distribution. While historical data will reflect the normal trends in inflows over the year (higher in spring and low in winter), it is important to note that these weekly inflow sequences are uncorrelated. Weather phenomena such as droughts, and the El Nino Southern Oscillation (ENSO), are particularly damaging for hydro management since they produce a sequence of low inflows, sometimes over months. Modelling the probability of a lower-than-expected inflow as being independent from week-to-week may produce less extreme results than are seen in reality.

This chapter is particularly interested in the mean and variance of profit in order to evaluate the hydro firm's exposure to risk. However, an examination of the intermediate outcomes such as generation, spot price, water values and storage will be made when necessary.

The simulation is run for different levels of contracts, for each firm. Specifically, contracts are varied between 30% and 110% of the expected competitive generation level for each firm, in steps of 10%. Under the parameters described above, the average competitive generation level for the two firms, over the year, is 1500MW each[41], and so these percentages are plotted in steps of 150MW, from 450MW to 1650MW. Since Scott's model was not concerned with either contract optimisation or a contract market, the strike price of these contracts does not have an impact on the optimality conditions within the single period model, or the reservoir management policy developed by the SDDP. The only contracting term of importance is the contract quantity of each firm (k in equation (8.3) above).

8.3.4 Interpretation of Output

The graphs in the following sections give summary statistics of the distribution of market outcomes, and are plotted for each combination of the two firms' contract levels (i.e., for each firm, from 450MW to 1650MW on contract).

Given that the model provides weekly profit figures, for every week of the year, over 40 different annual inflow sequences, there are a number of perspectives from which to view the data. Firstly, for each year, the variance of weekly profit, around the mean weekly profit, could be measured. We believe it would be incorrect to equate this measure of variability with "risk" or profit volatility. As argued in Chapter 2, not all of the variability observed in profit is "risk", since a portion of it is predictable. The way weekly inflows, and thus water values, release, and profit, vary over a year is often largely deterministic, in response to seasonal effects. Even if, in a particular year, inflows were deterministic, the variance of weekly profit over the year would be positive, reflecting the seasonal variations.

Secondly, we could look at the distribution of a particular week's profit, across the 40 inflow sequences, and compute the mean and variance. Averaged over the 52 weeks of a

---

[40] The inflows used were weekly, 52-period profiles obtained from 40-year historical data. The optimisation itself assumed (but did not require) that inflows follow a Normal distribution.

[41] This is using an expected marginal water value of $24/MWh, and averaged over summer and winter demand levels.

given year, this would give a valid estimate of how the variability in inflows, in a given week of the year, was reflected in the variability of profit in that week.

However, we believe that the firm is more likely to consider variations over a longer time period to reflect financial risk. A firm can potentially sustain high weekly profit variations, if over a long time period they cancel out. Hence the figures below show the mean profit over the year, for each inflow sequence, and the variance of the mean profit.

8.4 The Relationship Between Contracts and Profit Risk

Figure 8.5 and Figure 8.6 show the mean profit for Firm 1 and Firm 2, respectively, for different combinations of both their own and their rival's contracts[42]. Profit is decreasing in a firm's own contracts, reaching the competitive level of profit when both firms are contracted for the (expected) competitive generation quantity, and thus are generating at the competitive level.

[42] It should be noted at this point that a total and irreversible computer failure prevented us from completing this analysis. Additionally, the failure corrupted a number of data files, and the profit results plotted are from an earlier backup, which included an error in the calculation of the mean and standard deviation of the profit. The plots of all other variables (i.e., generation, spot price, storage etc) are accurate, but the contract revenue was omitted from the last 32 weeks of each year (and thus the calculations are dominated by spot profit only). Since the contract revenue was a fixed amount, for each contract combination, this does not distort the standard deviation of profit, since each sample point is distorted by the same amount, but simply increases the mean profit for each contract scenario (and the increase can be easily calculated). Clearly, the error is most significant at high contract levels, when the (g-k) term in the calculation of spot profit is smallest, however, this does not affect the trend in profit illustrated in the figures, i.e., profit is decreasing in contracts. The effect of this error on the estimation of risk will be discussed in Section 8.5.

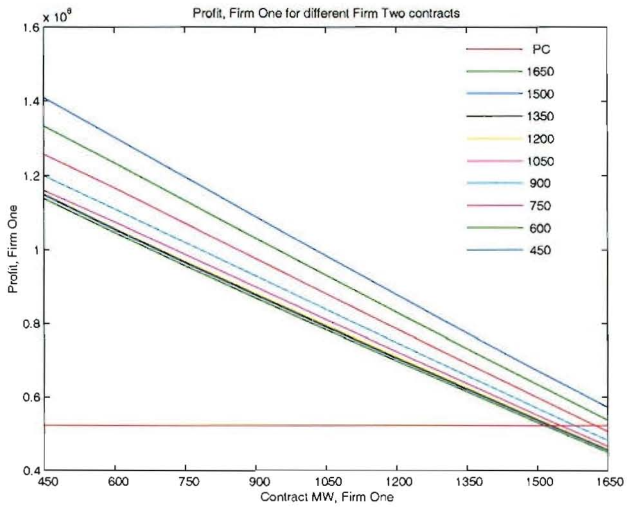


Figure 8.5 Profit, Firm 1 (Hydro) [43]

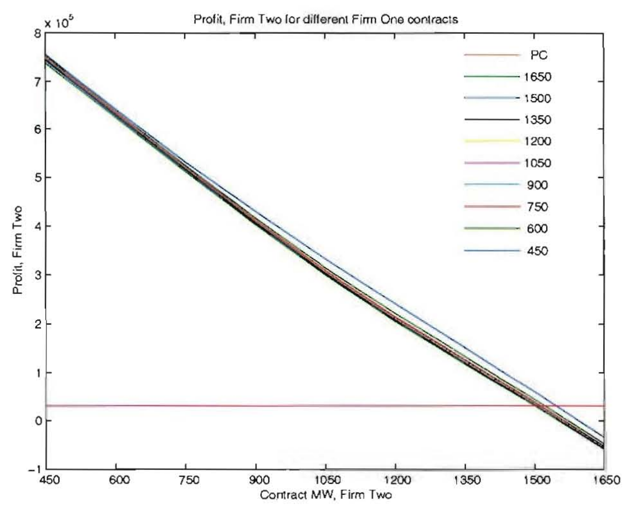


Figure 8.6 Profit, Firm 2 (Thermal)

[43] The figures can be interpreted as follows. Each coloured line corresponds to a particular fixed level of contracts, in MW, for the rival firm, as indicated by the key in the top right corner. The "PC" line, in red, indicates the profits (or, in later figures, the value of any variable) observed by the firm if it acted as a perfect competitor (i.e., if it equated marginal cost with price). As shown by Scott, and in later chapters here, a perfect competitor's generation output is not affected by its contract level, and hence the PC line is seen, in all figures, to be horizontal. Since Scott (amongst other authors) has shown that "full contracting", i.e., selling a quantity of contracts equal to the competitive output, leads to the firm behaving as a perfect competitor, we would expect to see results converge to the red line when both firms have sold 1500MW of contracts.

Since the optimal generation level exhibits an increasing relationship with contracts, the firms must be producing more, at a lower spot price, with a positive level of contracts, than they would do in the absence of contracts (ignoring capacity constraints). Under the assumptions employed here, these contracts are valued at the mean spot price for the year[44]. Hence as contracts increase, the effect on profit is a function of both the change in contract revenue, and the change in spot profits. Since we know that a firm's generation tends to the competitive quantity, as contracts approach the competitive quantity, the proportion of output that the firm (directly) receives the spot price for is decreasing. Since the spot price is also decreasing as output increases, and costs are increasing, we would expect that spot profits decrease at a greater-than-linear rate. However, contract revenue is likely to be increasing with higher levels of contracts, although the price at which contracts are valued is decreasing at the same time.

Analytically, we can define total profit as a function of the contract level by substituting optimal expressions (8.3) and (8.4) into profit function (8.2). Even assuming constant marginal costs, this yields a complicated expression, but after some manipulation in Maple, the derivative can be expressed as:

$$\frac{d\Pi_i^*}{dk_i} = \overline{p}(k_i) - \frac{2\rho^3\left(k_i + k_j - g_0\right) - 2\rho^2\left(2\pi_i\left(k_i\right) + \pi_j + p_0\right) - \pi_i\left(k_i\right)}{9} \tag{8.5}$$

The mean spot price, and marginal costs in this equation have been expressed as general functions of the contract level, since they are the result of a complex optimisation determining the water value surface and which step of the marginal cost curve the thermal firm operates on, thus determining the equilibrium quantities and prices in a given period. In any case, it can be seen from (8.5) that the function is not linear. However, the extent of the non-linearity is not great. The coefficient of the firm's contract level in (8.5), i.e., $2\rho^3/9$, is very small. For the parameters used here, this is

44 In reality, we would expect that the firm would receive a premium on the contracts that would be sufficient to induce them into contracting, compensating them for the loss in spot profits. This issue is addressed in the following chapters.

equal to 2.7 x 10^{-8}. Hence it is not surprising that the simulated contract-profit relationship, illustrated above, appears linear.

However, any interpretation of the average profit curves illustrated above must be made in the light of the assumption that contracts are valued at the mean spot price. If the contract price is determined by some other mechanism (e.g., consumer risk aversion), the additional contract profit received by selling more contracts may outweigh the loss in spot profit, and we could observe an increase in total profit with contracts, over some range of values. This issue is dealt with more fully in Chapters 10 and 11.

The variance of profit differs significantly between the two firms. The standard deviation of profit, for any level of contracts, is driven directly by three varying elements:

- Generation,
- Spot price, and
- Generation costs.

For the hydro firm, generation and short-run variable costs are determined by the water value implied by the level of inflows (via the reservoir management optimisation). Variability in these water values leads to changes in the position of the hydro firm's Cournot reaction function, thus affecting the thermal firm's equilibrium generation, and consequently the spot price. It is the combined effect, or co-variability, of these variables which ultimately determines profit variance. For a given contract quantity:

- An increase in the water value lowers profit maximising generation for the hydro firm (negative covariance),
- An increase in one firm's profit maximising generation results in a decrease in its competitor's profit maximising generation, capacity permitting (negative covariance, via the negatively-sloped reaction functions), and
- An increase in total generation decreases price (negative covariance).

Some of these effects are stabilising, while others amplify the underlying variance in inflows. The aggregate effect is difficult to determine analytically, since it depends heavily on the thermal firm's marginal cost steps. Numerical results are shown in Figure 8.7 and Figure 8.8.

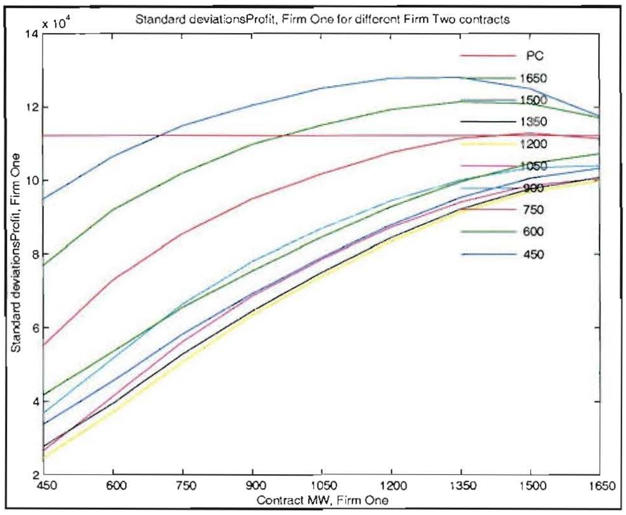


Figure 8.7 Standard deviation of profit, Hydro firm

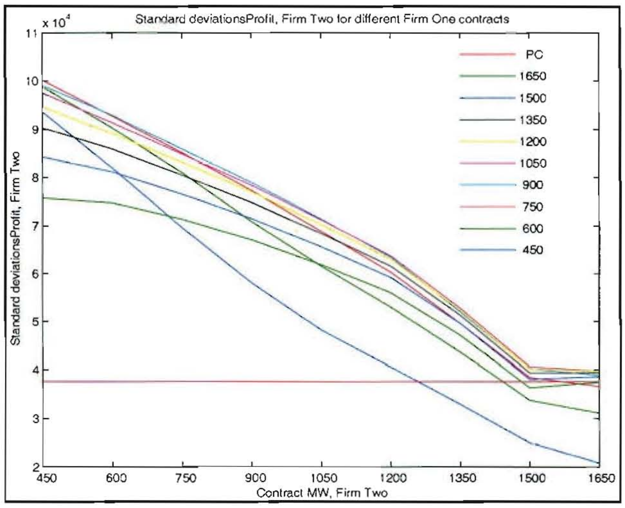


Figure 8.8 Standard deviation of profit, Thermal firm

The hydro firm's profit variability generally increases with the level of its own contracts. In the case where the thermal firm is not highly committed, the hydro firm's profit actually begins to decrease for levels of contracts above 1350MW. This reflects the relatively unconstrained responsiveness of the thermal firm to hydro variations, when it is not operating close to a step in its marginal cost curve (see below).

It is also clear that the variance of profit for the hydro firm is decreasing in the level of its rival's contracts.

The variability of the thermal firm's profit displays the exact opposite relationship with its own contracts, and is largely unaffected by the hydro firm's contracts. As its own contracts increase, the standard deviation of profit decreases, although this trend appears to flatten off at high levels of contracts.

The behaviour of these curves can be explained by looking at how the standard deviation of the three variables listed above changes with the level of contracts.

8.4.1 Generation, Firm 1 (Hydro)

At low levels of contracts, the firm is holding generation back from the market in order to increase the spot price and maximise profits (Figure 8.9). The low output of the hydro firm gives rise to high storage levels, implying that the firm will be operating on the flat section of the water value surface in most periods (see Figure 8.4 above). Variations in storage will give rise to only small changes in the water value, and thus we observe low levels of generation variance. Hence, instead of responding to varying inflows by altering its output, the firm prefers to absorb them in the reservoir and stabilise generation.

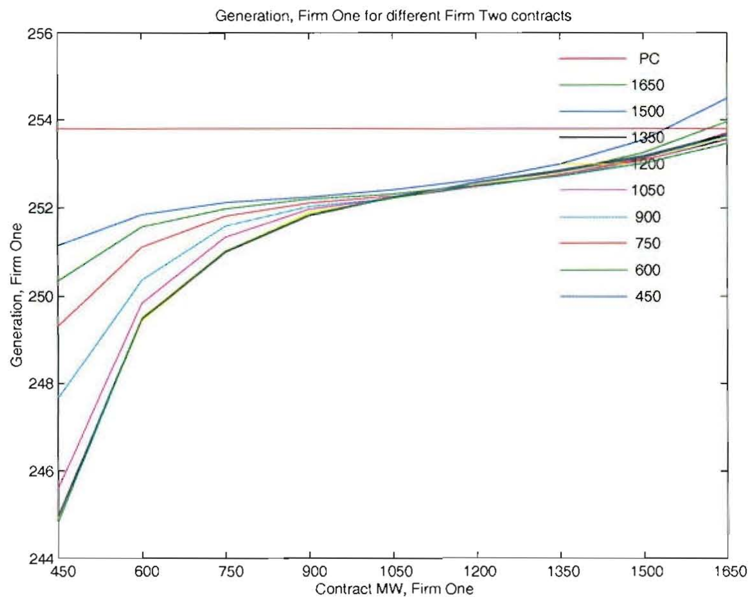


Figure 8.9 Mean generation, Firm 1 (Hydro)

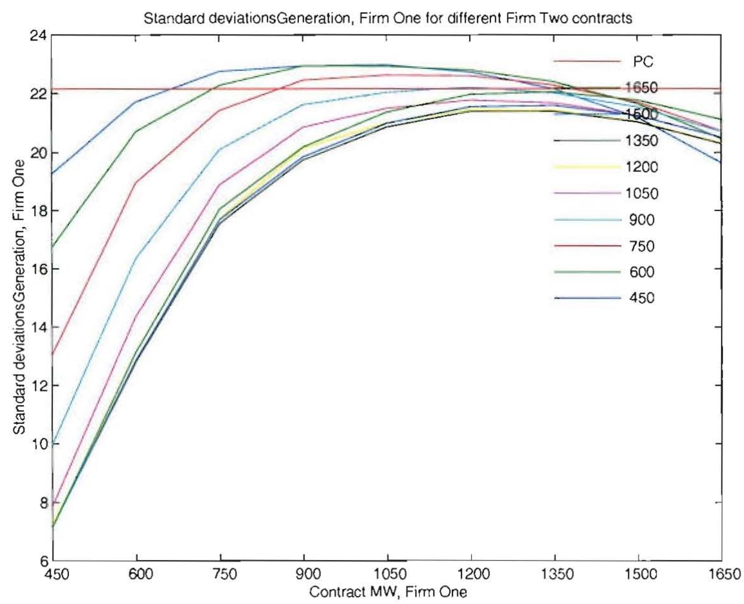


Figure 8.10 Standard deviation of generation, Firm 1 (Hydro)

As contracts increase from 450MW, the variance of Firm 1's generation initially increases (Figure 8.10). These moderate levels of contracts lead to higher mean generation levels and a decrease in the average amount of water in storage (Figure 8.11). Hence the firm, on average, is moving towards the steeper section of the WVS, indicating that in some periods (particularly winter), the marginal water value, and thus generation, may become quite sensitive to changes in the storage level.

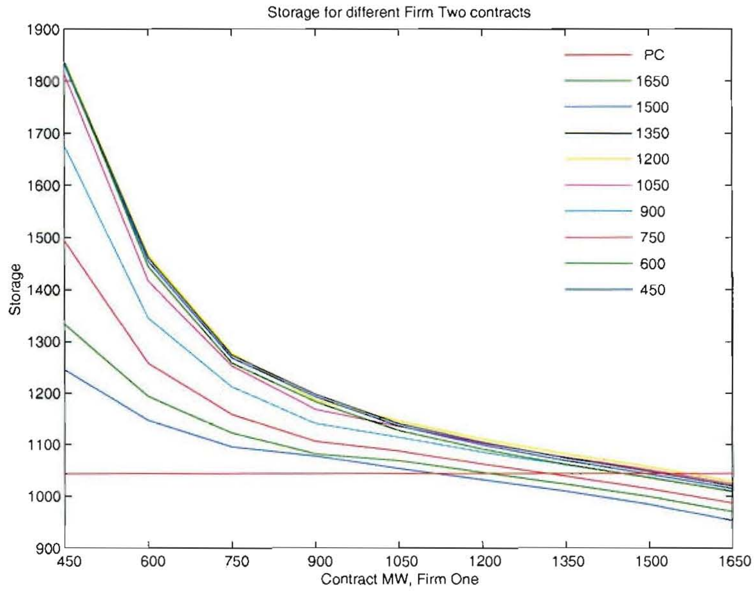


Figure 8.11 Storage, Firm 1 (Hydro)

As contracts reach 1200MW, the high levels of generation result in very low levels of storage. This implies that the firm will be operating on the steep section of the WVS, and thus marginal water values will be both high, and very sensitive to changes in storage, effectively penalising the firm for being close to "empty". Thus the firm responds to inflow variations by altering its generation in an attempt to stabilise the storage level (Figure 8.12). While the high water values provide incentives to conserve water by restricting output, the firm's high contractual commitments demand a high output. Thus in some situations, the firm may prefer to conserve water and "buy back" generation from the spot market in order to meet its contractual commitments. Such a situation would be observed when the marginal water value is high enough to make the fractional term in (8.3) negative, and thus its equilibrium generation less than its contract level. This is the highest-risk scenario for the firm discussed early on in the chapter, where the hydro firm buys back generation from the spot market, at spot prices that have been driven high by low storage levels and thus low hydro output.

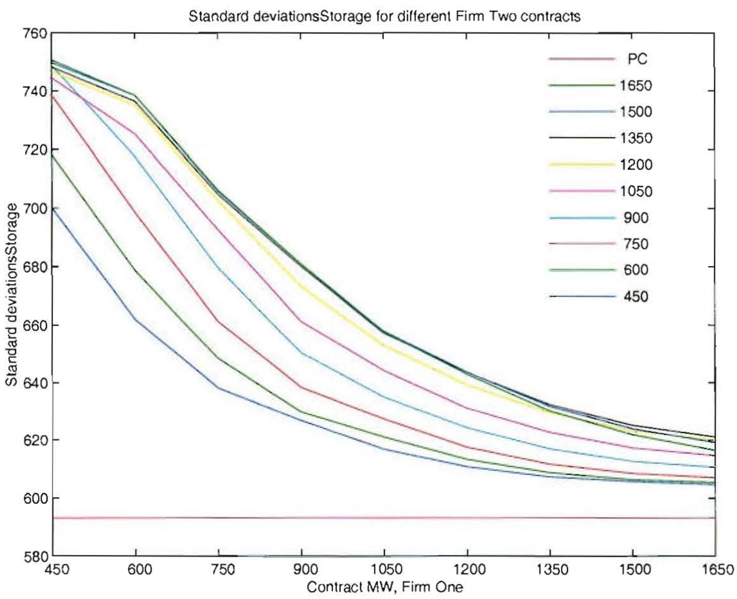


Figure 8.12 Standard deviation of storage

8.4.2 Generation, Firm 2 (Thermal)

At low levels of contracts, Figure 8.13 shows the thermal firm is operating, on average, somewhere in the interior of the "flat" corresponding to the \$30/MW section of its marginal cost curve (which applies to generation levels between 750MW and 1500MW, see Figure 8.2). Recall that the only varying factor in the market is the level of inflows received each period, and thus any variability in the thermal firm's output is in response to hydro output (or, more correctly, the marginal water value). At these low levels of contracts and generation, the thermal firm has a moderate amount of freedom to respond continuously, according to its profit maximising reaction function, to changes in the hydro firm's generation, since its marginal costs are constant for a reasonable range of output levels around the average level.

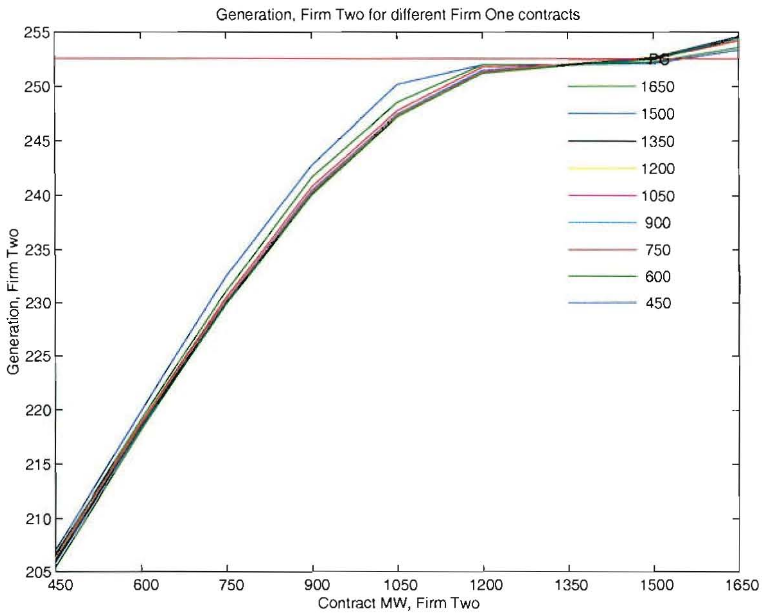


Figure 8.13 Mean generation, Firm 2

Hence we see thermal generation varying considerably at these contract levels, responding to any variations in the hydro output (Figure 8.14). It can be seen that this variance is greatest for moderate to high levels of hydro contracts, which, as shown above, result in the greatest degree of hydro variation.

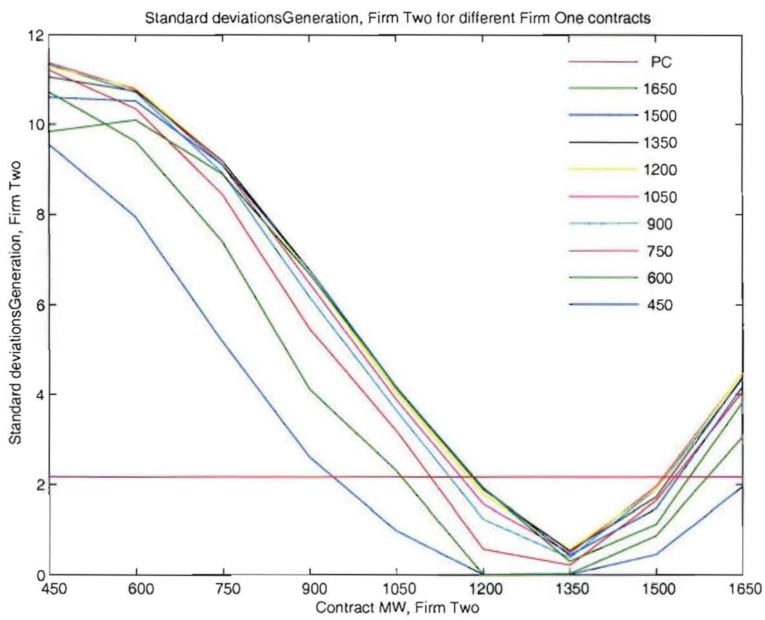


Figure 8.14 Standard deviation of generation, Firm 2

As the thermal firm's contracts increase, the thermal firm dramatically reduces its generation variability, to almost zero (implying that the firm is submitting an almost constant amount to the market) at 80-90% contracting (1350MW). There are two reasons for this. Firstly, as noted above, the firm is producing closer to its fixed contract quantity, and does not find varying from it profitable to any great extent. Secondly, the firm's typical generation quantity has increased with the contract level, thus pushing it closer to the step in its marginal cost curve, at 1500MW, corresponding to a sudden, and significant ($40) increase in marginal costs. The firm no longer responds continuously to hydro variations, as some of these responses would imply generation levels above the range of the $30/MW step. In particular, a generation level greater than 1500MW would attract a marginal cost of $70/MW. The thermal firm's output is almost constant at 1500MW, for contract levels between 1200MW and 1500MW. Only when the firm is over-contracted does it begin to operate on the next supply curve step, evidenced by the increase in generation variability at a contract level of 1650MW. Here the firm is trading off producing the extra commitments itself, at a high marginal cost, with effectively buying it back from the spot market, at spot prices.

8.4.3 Spot Price

As the firms respond to their contract levels and marginal costs, and "game" on the spot demand curve, the spot price varies also. Note that it is the variation of the *aggregate* quantity supplied by the firms that is reflected in the spot price variance, rather than each firm individually.

With the thermal firm offering a virtually constant quantity at high levels of contracts, we expect to see the hydro firm's variation reflected in spot price volatility at these firm 2 contract levels (black, blue and green lines in Figure 8.15).

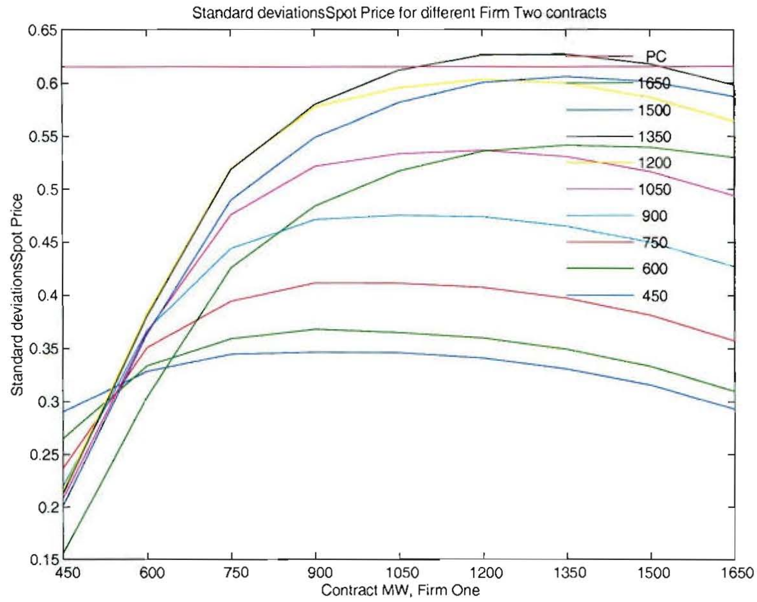


Figure 8.15 Standard deviation of spot price, as Hydro contracts vary

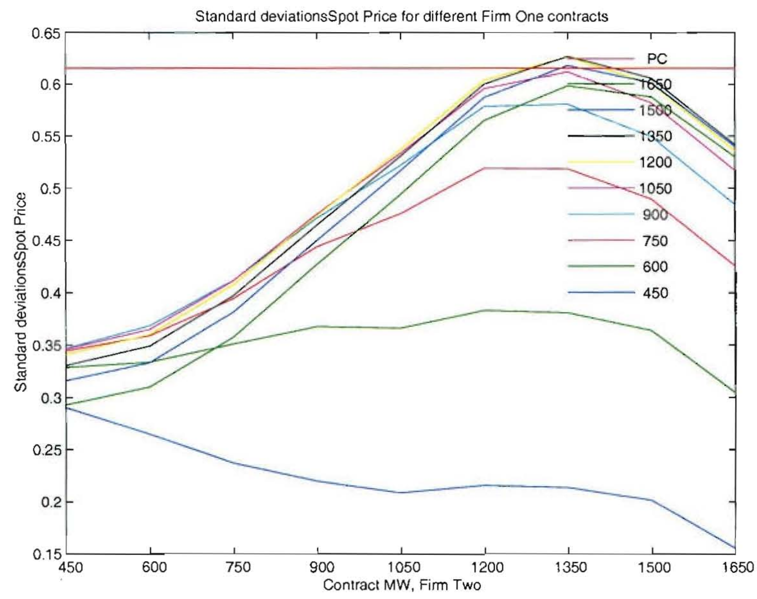


Figure 8.16 Standard deviation of spot price, as Thermal contracts vary

However, as the thermal firm's contracts decrease, its ability to profitably respond to variability in hydro output increases, and the spot price variance decreases (Figure 8.16).

The impact of the firms' combined output variability on spot price volatility is clear. Volatility is at its highest when both firms are highly contracted, since the hydro firm is attempting to stabilise (low) storage levels by varying its generation in response to inflows, and the thermal firm is choosing not to respond to these variations and is holding its generation constant. Spot price variance is generally at its lowest for low levels of thermal contracts, regardless of the hydro firm's contract level, since this is when the thermal firm can respond freely to hydro variations. However, it is noted in Figure 8.16 that the combination of contract levels resulting in the minimum spot price variance is for a high level of thermal contracts (hence stable thermal generation) and a low level of hydro contracts (hence stable hydro generation).

8.4.4 Summary

In order to understand the profit risk a firm is exposed to in a mixed hydro-thermal market, we have examined the behaviour of the underlying variables in a market where the only uncertainty is hydrological inflows. We have been able to explain the relationship between the contract level of the firms and the mean and standard deviation of storage, water values, generation, and spot prices for each firm, which are all key components of profit.

However, a number of the insights developed are likely to be specific to the particular industry structure analysed, and, in particular, the shape of the thermal marginal cost curve or the capacity of the hydro reservoir. These variables will have a significant impact on the shape of the water value surface, which reflects the value of hydro generation both in the current period, and in all future periods.

Notwithstanding this model-specific behaviour, we believe that some of the insights developed about the variability of firms' profits can be generalised to many markets that are dominated by hydro generation with only a moderate amount of reservoir capacity. In this situation, the optimal management of the reservoir will imply a general shape to the water value surface used by the hydro firm to determine profit maximising output levels[45]. The WVS will be steep close to the lower storage bound, reflecting incentives to conserve water by reducing generation. For moderate to high storage levels, the WVS will be flatter, although the exact shape will depend on the shape of the supply curve from the rest of the industry. The flatter section of the WVS indicates that the value of water does not vary significantly with changes in the storage level, as the firm is not close to either the upper or lower bound on storage, and that there is enough water in storage to imply that, in future periods, the spot price will be set by relatively cheap baseload, and potentially shoulder, generation This shape of the WVS gives rise to the following general market behaviour:

- At low levels of contracts, the hydro firm is holding water back from the market, since it is producing close to the low-output zero-contract Cournot solution. This gives rise to high storage levels, and stable water values. Thus the firm chooses to absorb inflow variations in the reservoir, and generation is relatively stable.

- As the hydro firm's contracts increase, its profit maximising output levels increase, and thus average storage decreases. This will bring the firm closer to the steeper section of the WVS, and water values will become more sensitive to changes in the storage level (i.e., inflow variations). These more volatile water values increase generation variability, as the firm gradually transfers from using storage to manage inflow variability, to using generation.

Hence low contract levels are characterised by low costs and generation, and low variability in both variables. High contract levels are characterised by both high levels and variability of generation and costs. Whether these trends are true for profit also depend on the effect the firm's behaviour has on the spot price. This, in turn, is a function of both the nature of the demand curve, and the response of other firms in the market. A monopolist may experience a degree of profit stabilisation from a downward sloping demand curve, since a reduction in output will induce an increase in the spot

---

[45] The general shape described below will be observed regardless of whether the firms behave a lá Cournot or not. However, as noted by Scott, a centrally coordinated competitive market would result in a WVS that more clearly reflected the thermal supply curve (i.e. piecewise linear step function, rather than downward sloping)

price. A small perfect competitor[46], on the other hand, will not experience this negative correlation between price and output, since the market price will not respond (to any great degree) to their output variability.

However, in the case of a two-firm Cournot market with smooth cost functions, firms generally react to increases in their rivals' production by decreasing their own production, and vice versa. This tends to stabilise total generation[47], and thus the spot price, and hence reduces the extent to which even dominant firms can experience the stabilising effect of a negative correlation between output and price. But the extent to which a firm is willing to respond to variability in their rival's output will depend on the impact their own response has on marginal cost. A firm with constant marginal costs can respond relatively "freely", while a firm with a more aggressively sloped marginal cost curve will be less inclined to increase production when its rival reduces output.

The impact of the shape of the marginal cost function was clearly displayed above. The particular positions of the marginal cost steps gave rise to the following observations:

- At low levels of contracts, the thermal firm is generating somewhere in the middle of a marginal cost flat, i.e., marginal cost is constant within a reasonable range of the average generation level. Here the thermal firm is relatively free to vary its generation in response to hydro variations.
- As contracts increase, profit maximising thermal output also increases, and thus the firm's average generation level moves closer to the marginal cost "step" at an output of 1500MW, from \$30 to \$70/MWh. Hence the firm may find its upside variations in output, in response to hydro volatility, restricted by this step, as the firm will prefer to stay at the end of the \$30 step, rather than move up on to the

46 Or any firm facing a very elastic demand function

47 It could be argued that if the response of the rival was large enough, it would actually *de*stabilise the effect of varying generation on the spot price. However, for this to occur, the rival would have to alter its generation, in response to the firm's actions, by an amount *larger* than the firm altered its output by in the first place. However, Scott (1998) showed that the Cournot game played in this model satisfies conditions set by Tirole (1992), guaranteeing unique equilibria. One of these conditions is that the reaction functions have slope (in absolute value terms) less than unity, implying that firms never respond to their rival's actions, by an amount greater than the rival's change in output.

$70 step. As contracts increase further, more of the optimal thermal responses to hydro variations "bunch" at the base of this step.

- For a contract level of around 1350MW, almost all thermal output levels are at the base of this step, implying almost zero variance in thermal output. Hence the thermal firm is not responding at all to hydro variations.
- For even higher levels of contracts, the thermal firm finds output levels corresponding to the $70 step profitable, and we observe it beginning to vary in response to the hydro firm again.

For a general thermal supply curve, this "bunching" of solutions will depend on the placement, and magnitude, of the marginal cost steps. Since the step solutions occur when the residual marginal revenue curve intersects the vertical section of the marginal cost curve, a smaller step in the above example would result in fewer solutions being located at the step, and thus the thermal firm would experience a greater freedom to respond to hydro variations. If more steps existed within the typical range of generation solutions, we may observe "bubbles" of thermal generation variance, where a thermal firm reduces its output variability as its average output moves close to a step, and then increases again as it moves onto the next flat section. As the steps move closer together, and get smaller in magnitude, these bubbles may eventually disappear[48].

Hence we observed high spot price variance when hydro variation was at its greatest (high levels of hydro contracts) and thermal response at it lowest (high levels of thermal contracts). Spot price variance was low both when neither firm was varying generation (low hydro contracts and high thermal contracts) and when the thermal firm had the greatest ability to profitably respond to hydro variations (low thermal contracts, any level of hydro contracts).

However, in general, we would expect to observe the greatest spot price variance whenever either firm is unable to offset its rival's changes in output in response to some

[48] In the limit, this would be equivalent to smooth, increasing marginal costs, which we argued above would still allow the firm to continuously respond to hydro variations, but not as freely as when marginal costs are constant

uncertainty that the rival firm faces. The least variance will be observed either when neither firm faces incentives to vary output (i.e., in response to uncertainty), or when at least one firm has the freedom to profitably respond to the other's output volatility.

We can now explain the trends observed earlier in profit variance for each firm (Figure 8.7 and Figure 8.8). At low levels of contracts, the thermal firm experiences constant, and thus stable, marginal costs within a reasonable range of its average output level. However, it is varying its generation significantly in response to hydro variations. Hence its profit variability is high. As contracts increase, it varies its generation less. It still faces somewhat stable costs, since it is yet to move up onto the next marginal cost step, and thus profit variance decreases.

However, the hydro firm faces it lowest profit variance at low levels of contracts. Marginal water values and generation levels are stable, as is thermal generation (since the thermal firm has little to respond to). As contracts increase, marginal water values and generation become more volatile, and profit variance increases. This is particularly evident for low levels of thermal contracts, which allow the thermal firm to respond, thus stabilising the spot price. Hence the correlation between the hydro firm's output and spot price is low. At high levels of contracts, the hydro firm's variance of profit is low when the thermal firm is relatively highly contracted, since this is when the thermal firm is generating a relatively constant amount, and the hydro firm experiences the undiluted effect of the negative price-output correlation.

Given the nature of the firms' profit maximising behaviour in response to inflow variability, we must now determine how significantly this impacts profit variance.

8.5 The Significance of Profit Risk

The simulation provided us with a distribution of the average weekly profit for each combination of the firms' contract levels, i.e., a statistical sample where each sample point was the mean weekly profit[49] over a year for a given annual inflow sequence. Each year of the simulation started independently of the previous year, so our sample points are not correlated in any way. As argued above, the mean and standard deviation of this distribution gives a reasonably accurate assessment of the unpredictable volatility in profit, arising from inflow uncertainty, which a firm will experience. This distribution of profit was chosen in preference to the distribution of profit over a year (i.e., between weeks) which would include a significant amount of predictable variation due to seasonal effects, and would also have correlated sample points.

If, for a variety of annual inflow sequences, the distribution of average weekly profit obtained does not have a significant level of variance, then we can reasonably conclude that the firm does not face a significant degree of risk resulting from inflow uncertainty. This section intends to estimate the riskiness of the resulting profit distribution, at each level of contracts, how the level of risk experienced by the firm varies with the level of contracts sold, and thus whether the opportunity exists for the firm to obtain significantly better risk positions, if we were to model the firm as being risk averse.

Assuming the firm followed mean-variance risk preferences, the utility it derives from each level of contracts represents a tradeoff between the long-run mean profit, and the variance of profit between years. A high level of variance may be tolerated in a situation that provides a high overall mean profit, while a similar utility will only be experienced in scenarios of low mean profit if the variance of profit is low. While we could postulate some risk aversion parameters and calculate the utility for each combination of contract levels, we have chosen the simpler option of expressing the standard deviation of average weekly profit as a proportion of the overall mean weekly profit, i.e., the coefficient of variation:

$$r_{k_H k_T} = \frac{\sigma_{k_H k_T}}{\mu_{k_H k_T}}$$

[49] It would possibly have been more intuitive to have used the total profit over the year, rather than the average. However, because of the technical difficulties explained in footnote 42, we only had access to mean weekly profit. However, this does not change the insights developed, since the average weekly profit is simply a scaled version of total profit over the year, i.e., total profit divided by 52.

where σ is the sample standard deviation, and μ the sample (i.e., overall) mean, of the mean weekly profits, in the scenario that the hydro firm holds k_H contracts and the thermal firm holds k_T contracts.

	30%	40%	50%	60%	70%	80%	90%	100%	110%
Hydro	0.04	0.04	0.05	0.05	0.06	0.07	0.08	0.09	0.11
Thermal	0.09	0.09	0.09	0.1	0.1	0.1	0.1	0.1	0.11

Table 8.2: *r* for different levels of firm's own contracts

For all combinations of contract levels, r was calculated for both firms to be between .05 and .1[50]. However, the relationship between the level of contracts and the level of risk experienced by each firm was significantly different. For the thermal firm, r was relatively constant at 0.1 regardless of the firm's level of contracts, or, in fact, the level of contracts sold by its rival (although r dropped to .09 for the lowest level of hydro contracts). This reflects the fact that contracts have a decreasing effect on both mean profit and the standard deviation of profit for the thermal firm, in a relatively proportional relationship.

On the other hand, the hydro firm experienced very low levels of risk (r = 0.04) at low levels of contracts, and much higher at high levels of contracts (r = 0.1). Again, r was relatively unaffected by the level of its rival's contracts.

Whether the coefficient of variation is an adequate measure of risk depends on the shape of the profit distribution. From the sample of 40 annual inflow sequences, mean profit

[50] As noted earlier (footnote 42), the profit figures, illustrated in Figure 8.5 and Figure 8.6, are in error, as contract revenue was omitted from the final 32 weeks of every year. Correcting the mean profit figures for this error would be a relatively simple exercise if the data files were still available (i.e., a constant amount, equal to average spot price multiplied by contract quantity, weighted by a the proportion of the year it was omitted for, would be added to mean profit). Our estimates put this error in mean profit at around 3% at low contract levels, and at high contract levels, the increase was 40% for the hydro firm, and as much as 100% for the thermal firm. Hence variance as a proportion of average profit, as presented in Table 8.2, significantly overstate the significance of the variance, especially in high-contracting states. This reinforces the conclusion that the profit risk was relatively insignificant.

was distributed approximately normally for each combination of contract levels[51]. While the standard deviation of profit was, at most, 10% of the overall mean profit level, for a normal distribution this implies that we can be 66% sure that the mean weekly profit, in any given year, will be within 10% of the long run mean profit level. At a 95% level of confidence, implying a z-value of 1.96, the width of our confidence interval is approximately 40% of the mean profit. While these confidence intervals are true for any level of contracts for the thermal firm, they are only appropriate for the hydro firm when it has sold 1500MW of contracts, and are much more favourable (from a risk-averse perspective) for lower contract levels. For example, at a contract level of 450MW, the confidence interval indicates that we can be 95% sure that the mean profit level in any given year will be within 8% of the global mean.

Given that the standard deviation of mean annual inflow, over the sample of inflow sequences used, was approximately 20% of the mean inflow, these results suggest that the combination of storage and market power on the hydro firm's risk-neutral, profit maximising decisions actually stabilises profit to a level that would be acceptable to a risk-averse decision maker, even though it does this unintentionally. The thermal firm also has good reason to be unconcerned about the variations in hydro output, as it too incidentally stabilises its profits as a result of its profit maximising decisions. Both firms will prefer low levels of contracts in this scenario, which values contracts at the mean spot price[52].

We believe these results suggest that either firm will find the level of risk they are exposed to satisfactory, even if they are not consciously operating in a risk averse manner. The samples obtained above suggest that both firms can be 95% confident that their profit levels in any given year will not be more than 20% below the long-run mean profit level. The hydro firm faces significantly less downside risk than if it adopts a low-

---

[51] However, it is worth noting that the distribution, while relatively symmetric, could also have been modelled as uniform. This would suggest that the measured standard deviation, with a Normal interpretation, potentially overstates the true spread of the distribution.

[52] It is worth noting that we also investigated the distribution of profit for each week, across the 40 inflow sequences. The standard deviation of inflows, in each week, was relatively constant over the year at approximately 75%, while the standard deviation of profit, in each week, varied between 10% and 20% of the mean profit. The contracts-standard deviation relationship was very similar to that described above. So while, within a week, the firm could experience moderate volatility in profit, the effect of market power and storage was to stabilise profits.

contract strategy. The thermal firm, on the other hand, is somewhat indifferent between contract levels using the coefficient of variation as a measure of risk.

8.6 Conclusions

Of interest in this chapter was the extent to which a profit maximising generation firm which faced uncertainty with respect to its fuel source would actually face financial risk, without deliberately altering its strategy to minimise it. The model developed by Scott allowed us to empirically investigate this, as it modelled two firms behaving as risk neutral, profit maximising duopolists, competing with Cournot conjectures, and incorporated optimal reservoir management by the hydro firm. Scott's model included the contract level as a fixed parameter, and hence by solving the model for various levels of contracts, we could examine the extent to which a firm, utilising its reservoir and market power in a profit maximising manner for an "inherited" level of contracts, was exposed to profit risk (defined as the standard deviation of profit) if it acted in a risk neutral fashion. If it was exposed to unreasonable levels of profit risk, then a risk-averse manager of a generation firm in a similar setting would do well to act in a way which deliberately reduced this risk. If not, then we can conclude that for the scenario modelled here, hydro uncertainty is adequately managed simply by maximising profits, and modelling the decision maker as risk averse, if he or she were indeed averse to risk, would not add much to the analysis. If the exposure to significant amounts of risk depended on the contract level, then a model of risk-averse reservoir and spot management (see e.g., Kerr, Read and Kaye (1997)), and including a contract optimisation, would be appropriate.

The key results of the chapter are:

- The relationship between contracts and profit risk was significantly different for each firm, and dependent on the characteristics of the thermal supply curve
- Profit risk is relatively small compared to the average profit earned by the firms

In order to understand the relationship between contracts and profit risk, we examined the behaviour of the firms in the market as contracts varied. The important insights were that the mean and variance of water values, and thus hydro generation, were low when hydro contracts were low, since the firm was holding water back from the market and maintaining a high level of storage. Hence the hydro firm's profit was high and stable. As contracts increased, mean generation levels increased, and storage decreased. This, in turn, increased the sensitivity of water values to changes in storage levels, and thus increased the volatility of generation decisions. Hence profit became more volatile as contracts approached the competitive output level. Since mean profit was also decreasing in contracts, high contract levels resulted in the greatest level of proportionate risk the firm was exposed to.

To a certain degree, the level of profit variability the hydro firm was exposed to depended on the behaviour of its rival, which, depending on the position and size of the steps in its marginal cost function, had a varying ability to respond to the hydro firm's variations, in such a way as to stabilise the spot price. Since we were modelling both firms as risk neutral, the thermal firm was free to engage in this responsive behaviour to the extent that it maximised profits, and this was observed to happen to the greatest degree when the firm was operating somewhere in the interior of a flat section of the marginal cost curve. In contrast to the hydro firm, the thermal firm experienced the greatest degree of profit variability at low levels of contracts, but this was largely because these levels of contracts caused its typical generation to be situated a long way from a step, on its marginal cost curve. As contracts increased, mean generation moved closer to the step (corresponding to a significant increase in marginal cost), and the firm's ability to profitably respond to any hydro variations diminished, and thus its profit was stabilised. It experienced its lowest degree of profit risk at 100% contracting. Like the hydro firm, average profit decreased with contracts, especially at levels that caused the firm to operate on a higher marginal cost step. Hence the thermal firm had a less obvious optimal contract position, since high profit also implied high risk, and vice versa.

If we were to increase the number of steps in the thermal supply curve, particularly around the typical range of generation levels, we expect that the same behaviour described above would be repeated. The extremes highlighted in the model (i.e., very high generation variance at low contracts, and almost zero generation variance at high contracts) were probably largely driven by the width of the "flat" on which the thermal firm operated most of the time, and the impact on costs of moving onto the next step, i.e., the height of the step. As the number of steps increased, we would expect to see "bubbles" of generation variance from the thermal firm, reaching a maximum when the contract level causes the firm to operate mostly on a flat, and a minimum when the contract level causes it to operate close to a step[53]. These bubbles would be reflected in the firm's profit, as it was above. For a high number of steps, the marginal cost curve would tend to a continuous, increasing function, and the bubbles would disappear. The firm would face increasing marginal costs, wherever it operated on the curve, and we would expect to see a lower degree of responsiveness to hydro variations, and thus lower profit variance for the thermal firm, regardless of the level of contracts.

It is also interesting to note that as the thermal firm responds less to hydro variations, the hydro firm's profit variance decreases. Due to a downward sloping demand curve, a reduction in hydro output attracts a higher spot price, and vice versa, and this negatively correlated relationship stabilises profit. In years of low inflows, or high contract levels (and thus highly volatile water values), if the thermal firm does not offset hydro variations, the hydro firm experiences this "natural hedge" as its output varies with changes in the water values. Hence a risk-averse hydro firm would rather its rival did not respond optimally to its variations. While no thermal response at all is largely unrealistic in a market such as the one modelled here (unless the firm is capacity constrained), the hydro firm would certainly observe a lower degree of profit variance itself if, as suggested above, the thermal firm had a continuously increasing supply curve, rather than the one with a large flat section, as used here.

Even though we can hypothesise as to the level of contracting a risk averse generation manager would choose, even the highest degree of profit volatility observed in the above model was not startlingly high, despite quite volatile inflows and only moderate storage capabilities. Over 40 simulated inflow sequences, the firms experienced mean profits that varied at most by ±20% of the long-run mean profit level, at a 95% level of confidence, regardless of the contract level. While the results clearly supported the

[53] Experiments with constant elasticity demand curves and various thermal supply curves clearly confirmed the existence of such bubbles in thermal profit, spot price, and, to a lesser extent, hydro profit, at the thermal contract

hypothesis that firms with market power preferred low levels of contracts to high levels, assuming the contract price is set by the mean spot price, we believe even risk averse hydro managers, and, indeed, their shareholders, would be comfortable with this level of risk. If contract prices were higher, then average profit would be greater (especially for high levels of contracts) and profit risk would be even less significant. Furthermore, firms did not consider the impact of their spot strategy on contract prices, which may have influenced their willingness to increase output at high levels of contracts, given that this would depress the spot price.

Linear demand curves potentially underestimate the market power of firms at high prices, and/or overestimate it at low prices. The choice of linear demand slope above achieved an elasticity of 0.25 at the competitive solution, but of course the elasticity would have a greater value at higher prices. The analysis above was also performed using constant elasticity curves[54], with elasticities ranging between 0.2 and 0.5, and even less risk was observed for low levels of contracts, and similar levels for "competitive contracting" (although the maximum level of risk was decreasing in elasticity, suggesting that an elasticity of 0.1, which some authors would argue is more realistic, would give greater profit risk). Since the demand curve was steeper, at higher prices, than the linear version, both firms experienced the natural hedge described above to a greater extent.

While the somewhat conservative assumptions employed with respect to the contract price (i.e., there was no risk premium, at any level of contracts) would suggest that we have produced an upper bound on the degree of profit risk, there are particular characteristics of the model used that might suggest otherwise:

- We expect that the size of the reservoir, and the quantity of inflows received, would have an impact on the hydro firm's ability to absorb inflow variations. The upper bound on storage represents a limit on the firm's ability to transfer water from times of surplus to times of shortage, since additional inflows when the reservoir is full must be spilled, despite the fact they could be valuable in future periods. If a firm spills frequently, an increased capacity to store water would

---

levels for which the profit maximising output was in the middle of a marginal cost flat.

increase average storage, which, as discussed above, could decrease both the average water value, and the volatility of water values. This would cause both an increase in profit for the hydro firm (from lower costs), and a decrease in profit risk. Of course, the reverse would be true for a lower storage capacity.

In the model used for this study, the reservoir reached its capacity in less than 8 weeks of the year, on average, suggesting that the effect, on the average quantity of water stored, from increasing the size of the reservoir, would be minimal for the same distribution of inflows used above. However, a lower capacity is likely to have a significant effect, increasing the mean and volatility of the water value, which would be detrimental to the firm's risk position. While this analysis was not performed, we believe that it is unlikely that a hydro firm with the degree of market share modelled here, in a hydro dominated market, would have a reservoir that had a significantly smaller capacity than 12 weeks of average inflows.

- A potential improvement to Scott's model would be to model correlated inflows. Recent events in New Zealand, outlined in the introduction to this thesis, have shown that the most significant periods of hydrological risk for a hydro firm occur when a prolonged drought is experienced, in which case inflows are highly correlated. While the distribution of inflows used in Scott's model did mirror seasonal effects, a low-inflow week could potentially be followed by a high-inflow week. This does not accurately depict the situation facing a firm during a season of drought. We expect that, while accounting for correlation may not change the average level of storage, it would introduce more extreme storage levels, thus increasing water value volatility. Whether the firms' combined use of market power, to maximise profit, would still result in relatively stable profit outcomes, requires further analysis.

While these modelling extensions may increase the level of profit risk the firm is exposed to, thus necessitating the inclusion of risk aversion (if indeed the decision makers are risk averse), the insights developed above suggest that developing a spot market, storage and

[54] Again, due to the technical difficulties experienced, we cannot display these results.

contract model under risk aversion would be extremely complex. For example, a traditional mean-variance model would not be very useful, since the relationship between contracts and risk, and thus the risk-return tradeoff, is potentially very complicated. Given that a realistic thermal supply curve may introduce a number of risk "bubbles", we believe only an integer model would adequately represent the relationship between contracts and risk, and this greatly increases the difficulty of finding optimal strategies.

In any case, we believe that the above results indicate that hydro managers who possess a significant amount of market power, and who act in a profit maximising manner, can generally be modelled as risk neutral without compromising the analysis greatly, even if they are in fact risk averse. Hence we will proceed to a model of a situation similar to the one above, but enhancing it to include a contract optimisation that allows the contract price to include risk premiums offered by risk averse consumers, and that will include the renegotiation of contracts given the spot behaviour of the generators. While input uncertainty for the generation firms will be included[55], we will assume that firms respond to this in a risk-neutral profit maximising manner.

We now turn our attention to an analysis of the problem faced by consumers of electricity.

[55] Although the management of hydro variations will not be as elaborate as the stochastic reservoir management model provided by Scott (1998)

9

DEMAND FOR CONTRACTS

9.1 Introduction

The established literature outlined in Chapter 7 showed that, for a dominant electricity supplier, selling hedge contracts resulted in higher output and lower prices in Cournot optimality. If the contract price is equal to the spot price (as is the case when risk-neutral speculators drive the contract price to the expected spot price), then selling hedge contracts, and responding in this fashion, results in lower profit for the generator. Under these conditions, a risk neutral generator would not find contracts attractive, and would not sell them (unless forced to by regulation). It was suggested that one reason dominant generators still sell forward contracts was that, in the absence of speculators, they receive a contract price greater than the (expected) spot price (known as normal backwardation), which offsets the loss in spot profit. Since, in many newly deregulated and/or relatively small electricity markets such as New Zealand, we observe very little liquidity (and thus speculators) in forward markets, this seems a reasonable rationale for dominant generators to sell CfDs.

In order to determine the spot-forward price spread, we must determine the size of the premium risk-averse consumers are willing to pay for electricity on forward contract, in order to avoid some or all of the risk they face. To find the optimal level of hedging for a dominant generator, we must also provide a model of how this contract price changes with:

- the quantity of hedges sold to consumers,
- spot market outcomes, and
- how many contracts the generator's competitors are selling.

This essentially requires us to model a market for electricity contracts, in a similar way to modelling a market for the electricity commodity itself (the spot market).

If we are to develop a contract market in this way, we must begin by modelling the demand for these contracts from consumers of electricity. This chapter provides an analysis of contract demand, assuming that contracts are well defined and relatively standardised, so that a continuous demand curve for contracts can be found. This demand curve will describe the relationship between the quantity of contracts purchased by consumers, and the price, per unit of contract, they are willing to pay for the implied level of hedging. This, in turn, will show the size of the premium consumers will pay to avoid risk.

9.2 Assumptions

The most accurate contract demand curve would reflect all issues relevant to the hedging behaviour of an individual type of consumer, across all types of consumers. However, many of the complexities surrounding decision making behaviour may not be conducive to a simple, analytical representation. Different modelling approaches mean different aspects can be modelled accurately, but at the cost of making general assumptions elsewhere. It is not the intention of this thesis to formalise all the issues surrounding optimal consumer hedging behaviour into a model. While this chapter attempts to capture some important aspects of consumers' attitude to contracting, ultimately it aims to find a convenient representation of general demand-side behaviour that is consistent with the purpose of the thesis. The various forms of demand-side contract and their implications for generator behaviour is the subject of another literature (see, for example, Gedra (1992) or Oren (2001)).

The following assumptions will define the boundaries of the analysis that follows:

- As discussed in Chapter 3, there are numerous issues that motivate electricity consumers to enter into forward contracts, such as investor security, financial risk, and the desire to discipline generator behaviour. However, we will assume that only the risk associated with electricity load uncertainty, and price uncertainty, is relevant to consumers' hedging decisions.

- As motivated in Chapter 2, we will initially use statistical variance as a measure of risk. Section 9.3 highlights some of the pertinent issues when modelling the process by which consumers form expectations and assess risk, and Section 9.5 continues this theme by presenting the mean-variance framework for risk averse decision making.

- The assumption of standardised contracts requires us to narrow down the types of consumers we can analyse. Many firms' procure electricity under elaborate pricing regimes and/or have complicated demand behaviour and hedging incentives. For this reason, we will only consider those customers whose hedging decisions can be conveniently represented by the fixed price, fixed quantity nature of these contracts.

The analysis presented also assumes that contract negotiations occur regularly but infrequently, for example once a year. Both generators and consumers aggregate all relevant information at hand to form expectations of costs and revenues for the coming year, and submit supply and demand functions for contracts to the contract "market". Once the contract level is set, the day-by-day electricity market operates as normal, given the contract level, until contract negotiations come round again. This is a reasonable approximation to reality in electricity markets such as New Zealand, and England and Wales (Powell (1993)). Section 9.5 provides an analytical framework for the consumer's contract decision making process.

Section 9.6 presents the important conclusions for the market equilibrium models of Chapters 10 and 11.

9.3 Consumers' assessment of risk

This section will discuss ways in which consumers might aggregate past observations of the electricity spot price, and transform it into expectations of future behaviour and an assessment of their risk. In essence, the former amounts to the construction of forecasting models, and how elaborate consumers choose to make them. This will be addressed in Section 9.3.1.

These models may include a number of other variables that help explain certain aspects of spot price behaviour. The complexity of the models that customers use is likely to depend on the potential for, and magnitude of, savings resulting from more accurate predictions. Small businesses with low electricity costs (relative to total profit) may have little incentive to invest in elaborate forecasting models, and will make somewhat "naïve" predictions of future spot price behaviour. Larger firms with significant electricity costs, who stand to gain much from accurately predicting the spot price for a given period (especially if they intend to vary their activity level in response to it) may invest in more complex models that incorporate some or all of the factors outlined below.

Later sections will show that for the types of consumers addressed in this thesis, profit risk is driven by one major factor: the volatility of the spot price. Section 9.3.2 reconciles the predictions made by the forecasting methods with our chosen measurement of risk or volatility, namely the statistical variance.

It should be noted that while some of the analysis that follows also allows for uncertainty in the electricity requirements faced by the firm (quantity risk), we will not discuss how firms form expectations of future load behaviour. This is omitted for two reasons. Firstly, many of the factors influencing load volatility are firm and industry specific, and beyond the scope of this study. Secondly, we will ultimately derive our aggregate demand curve for contracts based on the assumption that, while load may vary over the year, this variability is entirely deterministic.

9.3.1 Expectations of Future Spot Price Behaviour

At the simplest level, consumers may be considered to define the most recent spot price observation as the best predictor of the future spot price. More reasonably, when considering a contract of a certain duration, the mean spot price over the contract period would be more relevant to a consumer, and the predictor would be the mean for the most recent similar period, i.e.,

$$E(p_t) = \overline{p}_{t-1} \tag{9.1}$$

where $\overline{p}$ = the annual mean spot price, say. This prediction could be further embellished by using a greater history of past mean spot prices, as it seems unlikely that a consumer will only look at the most recent year for estimations of the spot price, unless there isn't any more information available (e.g., in the case where a new market has been established, or significant restructuring has recently taken place). For example, an auto-regressive analysis of past prices allows the consumer to evaluate the predictive information of the lagged observations. Using regression, a consumer could form a general model of expected price in period *t* as follows:

$$E(p_t) = \beta_0 + \beta_1\overline{p}_{t-1} + \beta_2\overline{p}_{t-2} + \ldots + \beta_n\overline{p}_{t-n} \quad n < t \tag{9.2}$$

where $\beta_0 \ldots . \beta_n$ are constants.

Customers may believe the predictive value of recent prices is greater than more distant ones. This would be reflected in the regression by the magnitude of the weights (β's) on each lagged variable.

Beyond this somewhat "naïve" prediction, consumers of electricity may be aware that the spot price is driven by a number of underlying factors, rather than just being a random variable. A spot price in a year may differ from other years not just due to random variation, but also because of certain conditions that prevailed in that year. These factors, discussed below, may be included explicitly in the forecasting model, or incorporated heuristically, as a method of "adjusting" the simple predictions outlined above.

Economic environment

Economic indicators, such as interest and inflation rates, will have a considerable impact on the operation of generation companies. An increase in the cost of finance increases a firm's total costs in the short run, and decreases the likelihood of capacity expansion and new entry in the long run. However, these aspects are beyond the scope of this study.

Threat of regulation

Lowrey (1997) used an empirical analysis of past spot market observations to show the effect of regulatory threats on the spot price in the British electricity spot market. It became very clear that the market behaviour of dominant supply firms became more competitive after threats by regulators to investigate issues surrounding market power. A change of government may, in fact, induce a similar effect, if supply firms suspect that the regulatory regime might change as a result.

If the market is dominated by large supply firms, it is reasonable to expect that these firms are withholding efficient production from the market and thus prices are being supported above marginal cost. The threat of regulation (or, indeed, independent entry) may induce these firms into producing a higher quantity, thus lowering prices, following the threat. While we will not include this factor explicitly in the model, we will assume that consumers will have interpreted past prices appropriately to account for this effect.

Knowledge of strategic behaviour by supply firms

Some consumers may consider it worthwhile developing strategic models of the market. Models such as those discussed in Chapter 4 and 7 could be built using estimated cost functions, which might be inferred if the consumer had access to information about the bidding behaviour of supply side participants. For example, consumers might believe that dominant supply firms compete with Cournot or Stackelberg conjectures, and hence a model based on those assumptions could be a reasonable method of predicting spot prices. Such analyses have been performed by, for example, Wolfram (1999), and critiqued by comparing the estimations of the theoretical models against observed spot prices. Such models are also useful if the market is about to change in structure (e.g., entry of cheaper plant).

However, this thesis assumes that the current market structure has not changed within the period covered by the consumers' spot price observations, nor is it expected to change in the near future. Hence knowledge of how the strategic interaction between firms helps determine these spot prices would not add much to the consumers' forecasting process.

Hydrological information

When hydrological aspects are considered, consumers' expectations of future spot prices will be influenced by whether the market is facing a "wet" year or a "dry" year, with respect to the level of inflows expected. Such issues have a significant impact on spot prices for electricity where hydro generation plays a large part. Recent examples of note are the New Zealand crises of 1992 and 2001. The New Zealand electricity market is dominated by hydro, and hence the effect on spot prices in the winter, when demand is at its peak, is quite noticeable.

If it is possible to classify each year as being wet or dry, or somewhere in between, then consumers can use such information in forecasting models. Consumers may heuristically or analytically adjust past data to incorporate the hydro state of the year. A recent observation of the mean spot price that is high does not necessarily imply that spot prices are beginning a long-term upwards trend – if that year was classified as dry, we can discount the effect of that mean in our prediction model, thus "normalising" each observation. Such a classification scheme could be determined by elaborate models using meteorological data.

Prediction of spot prices by hydrological conditions is more difficult, although in a hydro-dominant country such as New Zealand, the media efficiently disseminates meteorological predictions about coming critical periods, e.g., winter. However, contracts may have to be purchased well in advance of a time when certainty over the hydrological 'state' of the period is reached. For example, a large NZ retailer neglected to sign contracts in February 2001, for the winter (June-August). It was speculated that this was largely due to there being no indications of a dry winter, at the stage that the contracts were being finalised, and hence the fixed price contracts being offered to them (which were "cheap" in retrospect) seemed unattractive.

As argued above, unless the meteorological conditions are expected to depart significantly from recent patterns, knowledge of the cause of certain observations will not add anything to a model of expected spot prices in the framework presented here. Under the assumption that long-term hydrological patterns are stable, a large enough sample of past data will provide an effective measurement of the expected future spot price behaviour, although extreme conditions (say, 1 in 100 meteorological patterns, such as severe droughts) may be excluded.

9.3.2 Measurement of Risk

Once consumers have determined how they will interpret spot price observations in terms of their underlying factors, they must form the distribution of prices that will be used to estimate the risk that they are exposed to. As noted already, this thesis will use the statistical variance as the measure of risk, and a mean-variance function of profit as a model of decision making under risk aversion. However, exactly how the distribution is formed, and the variance calculated, determines how accurately the true riskiness of a firm's position is represented by this method.

As emphasised in Chapter 2, only part of the overall variability of spot price observation may present a risk to the firm. Spot price behaviour over the day, week, and even year can, to a certain degree, be explained and predicted based on day/night, week/weekend and seasonal effects, respectively. If the firm is aware of these patterns in advance, the variability does not present a risk. However, if the spot price varies beyond this established pattern, the firm is exposed to risk. We have previously called this unpredictable spot price behaviour "volatility". The overall variability of the spot price, over a given time period, is thus made up of both predictable and unpredictable elements[56], and it is the latter that we wish to represent as risk. Moreover, we must ensure that our chosen measure of risk, namely variance, is accurately reflecting this volatility (or at least as well as possible), rather than simply reflecting total variation.

---

[56] Which would imply that there exists an additional risk-hedging mechanism to the firm, i.e., improving forecasting methods so that less of the overall spot price variability is volatility and thus risk. However, this aspect of risk management is ignored in this thesis.

Load duration curves (LDCs) and price duration curves (PDCs) are helpful ways of describing the natural patterns of load and the electricity spot price over a given period. Based on past data, these curves illustrate the relationship between a given value of the load (spot price), and the number of hours that the load (spot price) has exceeded that value, and hence are very similar to cumulative probability distribution functions (Figure 9.1)[57].

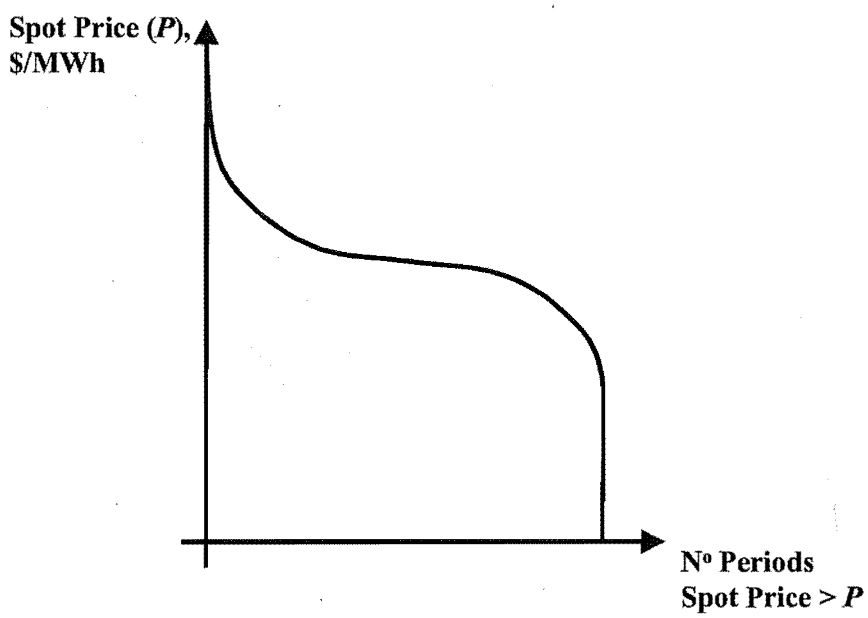


Figure 9.1 Price Duration Curve

By aggregating all past data into a LDC/PDC, the predictable patterns of spot price and load behaviour are reflected in a single function. While a PDC, for example, does not show exactly in which periods the spot price will exceed a certain value, it does illustrate the number of periods within which the firm will be exposed to such prices. If this information was known, the firm could prepare itself well in advance, as the price and load patterns are essentially deterministic.

[57] However, care should be taken in the interpretation of these curves. In this discussion, a LDC/PDC represents the range of individual spot prices observed over, say, a year, rather than a cumulative distribution of possible prices within a given half hour.

At a more realistic level, a firm may be certain about the range of prices that will be observed in a year, and their associated frequencies (or probabilities), but is unsure as to when the prices will occur. Here, the short-run responsiveness of the firm is critical - it anticipates the observation of the price, but does not anticipate when.

More significant risk arises, however, if particular loads or prices occur more, or less, frequently than the curve suggests. Such volatility is equivalent to uncertainty about the exact shape of the curve.

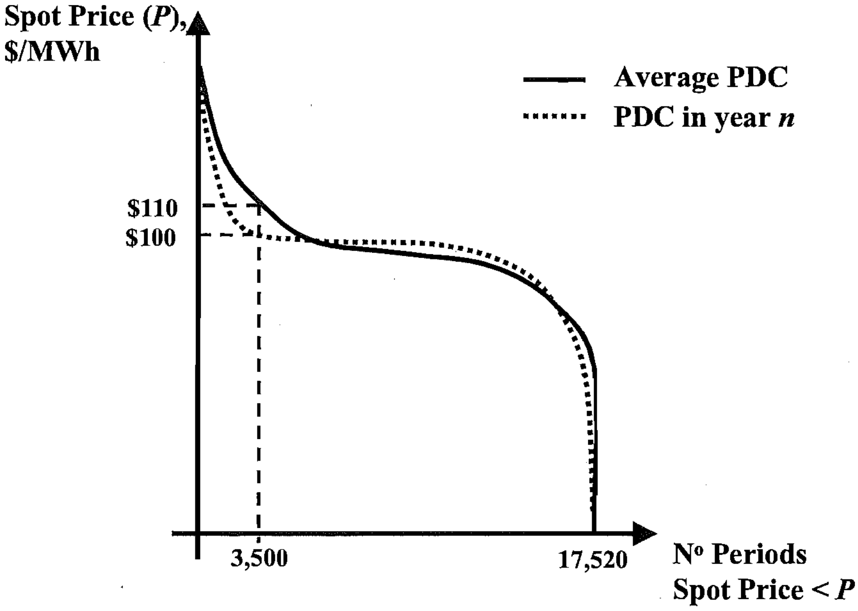
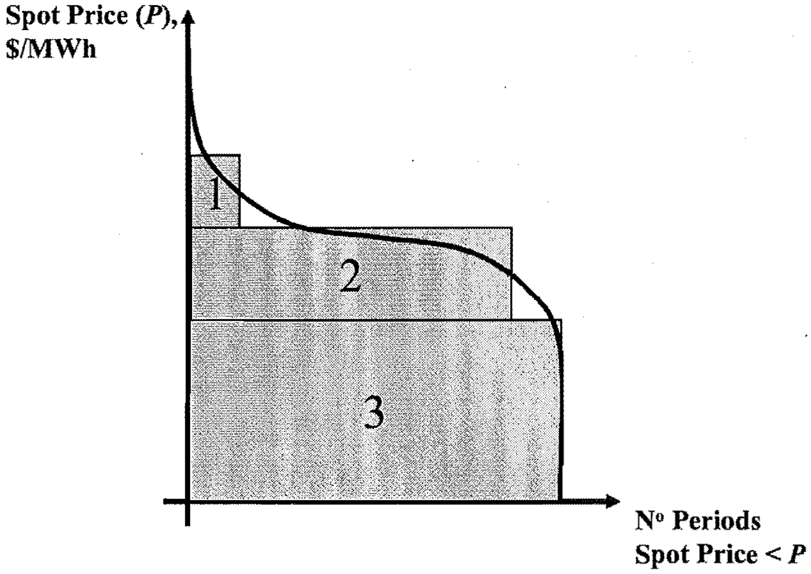


Figure 9.2 Uncertainty in a PDC

Figure 9.2 illustrates two annual PDCs, thus consisting of 17,520 half-hourly periods. The solid PDC represents the average price that is exceeded for each number of cumulative hours over all past observations, while the dashed line is a PDC observed in a particular year. On average, the spot price is expected to be at least $110/MWh in 3,500 periods. However, in a past year, it has been greater than $100/MWh in the same number of periods. In reality, there could be a number of possible prices that are exceeded in 3,500 periods, and the variance of these prices would reflect the volatility at this point. The variances at each cumulative period could be combined to form an overall picture of the uncertainty surrounding the shape of the curve, i.e., the volatility.

However, such a calculation would be complicated. In order to simplify the information contained in the LDC/PDC, it can be divided into a small number of sub-periods, each approximating the complete curve



(Figure 9.3). This is equivalent to dividing up the year into periods within which the spot price is in a certain range, distinct from the ranges defined by the other sub-periods (Figure 9.4). Risk is then measured as the variance of the average spot price (and/or load) observed within each subperiod in past years, from the overall average obtained from the expected curve.

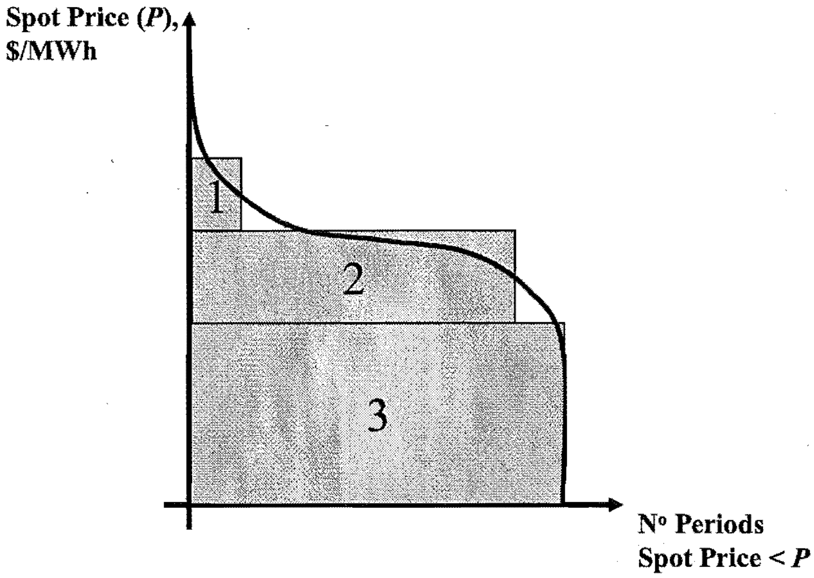


Figure 9.3 Approximation of PDC by 3 subperiods

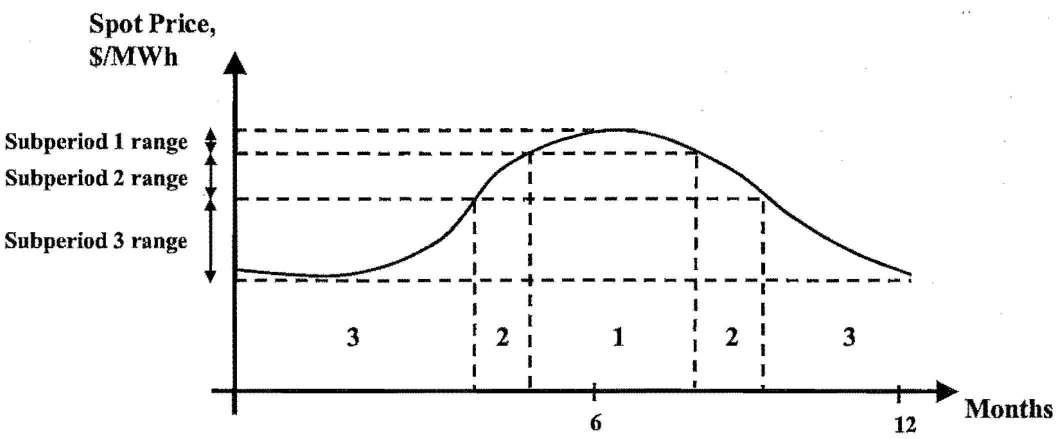


Figure 9.4 Prices in 3 PDC subperiods

While this method still loses some of the predictive ability of the LDC/PDC by aggregating across observations within a subperiod, it is superior to either taking the statistical variance of all half-hourly price and load observations, or, at the other extreme, taking a long-run average of these variables, and representing risk by the variance around it.

9.3.3 Assumed Model of Expectations and Risk

As outlined in later chapters, we will assume that consumers have access to a wide range of spot price observations from the past, and from these form a distribution of spot price behaviour. While it is tempting to include some of the additional factors outlined above (e.g. hydrological information), it is, in fact, irrelevant, since our model requires that expectations must be formed prior to contract negotiations and the period for which the contract covers. The implicit assumption is made that neither consumers, nor generators, know anything about the states of any of these variables at that point[58]. Hence consumers' best estimate of the spot price for the contract period is the average of all those previously observed. For example, a consumer's expectation of the coming year's spot price would not be changed if the most recent observation was high. The consumer would be aware that it was not the start of an upward trend in prices (unlike the "naïve" forecaster), rather, it was an observation that occurs with a given frequency, in their distribution of prices.

The implication is that consumers place no more "weight" on recent observations than they do on those in the more distant past. Instead, the weight consumers give a particular spot price observation is determined by how frequently it is observed.

If the spot price and/or load varies over time, we assume that consumers have sub-divided a PDC or LDC as described above, and the variance of the relevant variable is measured for each sub-period. Given that the models that follow solve for a single optimal level of contracting, one of two possible assumptions must hold:

- A different level of contracts can be signed for each sub-period. Since we do not model any correlation between sub-periods, a demand curve for contracts (DCC) can be formed, and the models that are presented in the following chapters can be solved for each sub-period of the LDC or PDC. If a firm is exposed to variability in **both** spot price and load, it is important that the sub-periods in each curve

[58] The model could be enhanced by allowing consumers to update their expectations as information comes to hand, using, for example, a Bayesian model. However, this is well outside the scope, and does not add much to the intentions of this thesis.

correspond to the same individual periods in the year. This ensures that the combined effect of price and load, on overall risk, can be accurately reflected

- Consumers have aggregated the PDC or LDC into one subperiod.

It is not important, to the analysis that follows, which of these assumptions is made. However, it should be noted that neither of the above assumptions is equivalent to assuming that a single level of contracting applies to multiple subperiods. That would require us to aggregate each of the variance measures taken in each subperiod, to provide an overall assessment of risk.

9.4 Demand behaviour for different consumer sub-groups

One popular approach to studying the effect of quantity and price uncertainty on hedging behaviour is to assume that a decision maker has an uncertain endowment of a commodity that they wish to sell. While the sale itself may take place in a spot market, the decision maker can hedge the uncertainty of both the spot price and endowment quantity with forwards contracts. In this sense, both price and quantity are random variables that are determined independently.

This framework is appropriate for modelling the decision process of some electricity consumers, but does not capture some significant characteristics of others. The discussion below presents some of the important and distinct characteristics of electricity consumers, so that hedging behaviour can be defined for each.

The decision makers, in this chapter, are those consumers who choose between direct exposure to the spot price and purchasing hedge contracts from generators for some or all of their electricity requirements. This includes Retail Electricity Companies (REC's, or retailers), who are intermediate sellers of electricity. The hedging behaviour of consumers or firms who purchase electricity **from** REC's will not be examined *per se*, but in aggregate they determine how REC's behave in the spot and contract market

However, we will not consider retail companies who are incorporated into a vertically integrated (VI) firm. While it could be argued that the relationship between a generator and a retailer in a VI structure is a form of a forward contract, there are a number of other incentives acting on the participants. Vertical integration may be construed as entry deterrence or price discrimination, and the form of the "contract" has different effects on the firm's risk position[59]. There is considerable debate in the industry and literature over whether VI is equivalent to forward contracts, and we do not wish to add to it here. In any case, the discussion involves aspects that do not make up part of this study.

9.4.1 Consumers with Unresponsive Loads

Many firms are either unable to, or choose not to respond to electricity spot prices by altering their electricity requirements. For example, a retail store for whom electricity is an overhead, and cannot, or chooses not to, adjust its trading hours in response to the electricity spot price, may fall into this category. Another reason that a firm may choose not to vary their electricity requirements in response to the electricity price is that electricity cost is such a small component of the total operating cost to the firm, that it does not warrant doing so.

Ignoring the possibility that the firm might attempt to conserve electricity in response to extremely high prices, we can assume that electricity load in this case is largely independent of the spot price. Some unresponsive loads may be largely predictable over a given time period, either not varying at all, or varying deterministically over time. The only risk faced by such a firm would be volatility in the unit electricity cost, which would be the spot price if it was purchasing directly off the spot market. Other firms, however, may be uncertain about both their load and the spot price in a given time period. Depending on the level of correlation between load and price (see below), this may imply a greater risk for the firm.

Retail Electricity Companies

In countries such as New Zealand, the vast majority of households and small-medium sized businesses purchase electricity through a Retail Electricity Company, or REC.

[59] While "locking" customers in may decrease risk, as contracts need to be re-negotiated each year, customers cannot be easily traded if the firm finds itself in an undesirable position in the contract market.

Retail firms purchase electricity at the wholesale spot price, and sell it to end-users under a variety of pricing arrangements which, in aggregate, provide a profitable margin for the REC. The supply agreements between a REC and its customer base are wide and varied in form, but are presumably constructed to match the position of the REC in the wholesale electricity market. Since only a portion of the REC's customer base will "see", and possibly respond to, wholesale spot price variations, it would initially appear that the amount of dependency between the aggregate load variation and spot variance would be small, making the traditional framework, outlined above, appropriate.

However, there do exist some significant relationships between load and price movements for retailers, and how predictable these relationships are will determine both the level of estimated correlation between these variables, and the risk implied by their variability. Day-night wholesale price differentials arise because of the changes in aggregate wholesale demand between these two periods. If an individual REC's customer load is highly correlated with system demand (which it will be if it has a representative portfolio of customers), its own load will exhibit a similar correlation with price. Since higher demand implies higher prices, the REC faces the risk of having to purchase a large quantity of electricity at a high spot price. The common REC tariff that incorporates day/night differentials is a mechanism that manages this risk, on a short time scale. Such a scheme provides incentives for consumers to conserve load in high price periods, thus reducing the retailer's exposure to the high wholesale prices.

On a medium-term scale, most countries experience seasonal variations in electricity demand. In warm climates (e.g., California or Australia) air-conditioning requirements increase electricity load dramatically in summer. In many other countries, including New Zealand, a similar winter effect is noted, as heating and lighting demand inflates aggregate load. New Zealand is a temperate and mountainous country, and electricity generation is dominated by hydro production. In winter months, some precipitation is "locked up" in the snowpack rather than being transformed into reservoir inflows. This results in lower reservoir levels, and a higher value being placed on water during these months. As a result spot prices are driven up (see Chapters 6 and 8). Here, the retailers' summer-winter variations in load are likely to be correlated with the spot price. However, this does not necessarily imply a risky position for the firm. Given that these seasonal effects are largely predictable, an REC may have them built into its tariff structure, and can make arrangements to deal with cash flow issues.

Risk does arise, however, when on the spot price and/or aggregate load vary away from the predicted pattern. To illustrate, in the New Zealand electricity market in 2001, a large REC decided against purchasing forward contracts for the winter period, perhaps because it believed that the contract price being offered was greater than the expected spot price for the period, even accounting for the seasonal effect that normally drove up prices during winter. As winter approached, many of the hydro reservoirs received significantly less water than normal. Wholesale electricity prices, reflecting the water shortage, were driven to four times their normal level, and more importantly, significantly higher than the price that had been offered on a hedge contract earlier. The REC was forced to purchase electricity for their customers at these high spot prices, possibly 2 or 3 times as much as they were receiving for it. This crippled the firm, and it eventually exited the electricity market as a direct result of this incident. While the forward contracts offered seemed unprofitable given the REC's prediction about the period (see Section 9.3.1), these contracts would have prevented the dire financial situation the firm found itself in.

Also contributing to quantity uncertainty for a REC is that it must compete with other RECs for its customers, as discussed in Chapter 2. Competition between retailers for consumers' electricity load, combined with a relatively fluid "switching" process, makes this more of a medium-term issue than a long-term one, and leads to new complications for contracting. To be insulated against spot variations through long-term contracts is potentially costly, since a REC could get locked into high contract prices while its competitors take advantage of low spot prices, thus attracting customers away from them.

For the purposes of this study, we will assume that all firms whose electricity requirements are unresponsive to the electricity price purchase their electricity through a REC. Each REC faces, in aggregate, a potentially uncertain and unresponsive load that it must purchase through the spot and/or contract market. Assuming the RECs are all risk averse, the optimal electricity purchasing decision for them is determined by maximising the expected utility of total electricity costs – which, under the mean variance model (see

Section 9.5.1), is equivalent to trading off the cost and variance of the cost of purchasing their electricity requirements.

9.4.2 Consumers with Responsive Loads

For many firms, electricity requirements are significantly influenced by their level of activity, and so the price of electricity becomes a variable cost of production. Hence the profit maximising output level, and thus their electricity load, could be expressed as a function of the unit electricity cost, defining an optimal "response" curve to the electricity price. For some firms, however, the sheer logistics and/or cost of changing their level of activity within the time-frame required makes such a response infeasible or unprofitable (when compared with no response). This would be true of a firm which, for example, faces a constant demand for a non-storable product[60], and thus electricity requirements are constrained by demand. Such firms would be considered as part of the group of electricity consumers, discussed in Section 9.4.1, that have relatively fixed loads, with respect to the spot price. Similarly, it is possible that only a small number of firms have electricity as a significant enough input cost, or are sufficiently risk averse, to warrant adjusting their production level in response to the spot price. Changes in a relatively insignificant electricity cost may only induce minor variations in profit maximising output, and the firm may prefer to keep production level fixed, or at least independent of the unit electricity cost. These firms would again be included in the group of consumers described above.

However, there are firms that may find it profitable to respond in this fashion, and have the ability to do so. The total cost of electricity purchases is usually significant for such firm's, and the potential gains from an optimal load response may be attractive. As an example, Jones (2000) developed a dynamic programming framework for a large cement company to decide how to schedule an energy intensive process (the operation of a kiln), in response to day-ahead predictions of the electricity spot price.

[60] If the price is not too high

This dynamic response may represent an additional form of risk hedging for such a firm. While the number of firms that are large enough to warrant load response is small, the firms themselves are often very large, and thus represent a significant proportion of total electricity load. If the response has an impact on their exposure to financial risk, we should certainly attempt to include their risk-averse hedging behaviour in our model of contract demand.

Responding to the electricity spot price variations does not necessarily imply the firm is attempting to hedge risk *per se*. Rather, it is *uncertainty* about spot price movements that determines the riskiness of the firm's position. If spot price movements follow a deterministic path over time, a firm would be able to schedule production runs well in advance in the manner described above. However, a firm that has the ability to respond may also do this in response to spot price behaviour that was unexpected, thus providing another level of risk hedging (as long as the response reduces profit variance). The work of Jones incorporated what knowledge the firm had of medium-term predictable variations in the spot price (for example, day/night differentials and seasonal effects) into the framework, and then used the firm's ability to respond at short notice by using the day-ahead price as a predictor of the spot price they would actually face. The sheer size of the firm's electricity load meant that even small unpredictable movements of the spot price away from its expected path meant significant swings in cashflow for the firm. The ability of the firm to respond at short notice provided an additional hedge against this spot price volatility, and thus would affect its incentives to purchase fixed price forward contracts.

9.4.3 Summary

Due to the complex array of pricing options offered by RECs to their consumers to attempt to manage the load-price correlation risk, it is difficult to model their problem explicitly. The REC decision problem is further complicated by the Bertrand-style competition between such companies in the new deregulated environment. In this case, we will assume that an individual REC's aggregate requirement is uncertain, and largely unresponsive to the spot price. Given that the vast majority of individual firms or households with unresponsive loads purchase through these retail firms, the REC will form the first major type of electricity consumer who desires to hedge, that will be modelled here (Section 9.5.3).

Additionally, we will address the problem of the large firm which purchases directly off the spot market and has the ability to respond to the electricity price (9.5.4).

9.5 Demand Curves for Contracts

9.5.1 Utility Functions for Risk-averse Decision Making

Utility functions are a commonly used method of representing an investor's attitude to wealth and risk. John von Neumann and Oskar Morgenstern developed a general class of utility functions, that give a utility, $u(x)$ for each wealth level x. General classes of utility functions, including von Neumann-Morgenstern (vN-M) utility functions are reviewed in Appendix A.

A special form of v-NM function is the mean-variance (M-V) function. M-V functions are popular in the financial literature both for their simplicity and intuitive appeal. Optima of M-V functions can be shown to be the same solutions as would be found if the decision maker maximised the expected value of a quadratic utility function, or, in the case where wealth is normally distributed, a logarithmic utility function (Levy and Markowitz (1979)). Decision makers who act according to M-V preferences are believed to exhibit a mix of expected-wealth maximisation, and wealth-variance minimisation. The exact mix is determined by a single risk aversion parameter. More explicitly, the expected value of the utility is defined as:

$$E[U(\tilde{x})] = E[\tilde{x}] - \frac{\lambda}{2} Var[\tilde{x}] \tag{9.3}$$

where $E[\tilde{x}]$ is mean or expected wealth, $Var[\tilde{x}]$ the variance of wealth and λ is known as the coefficient of absolute risk aversion, reflecting the investors attitude to risk or variance. A risk neutral individual would have $\lambda = 0$, implying that utility depended on expected wealth alone, while a $\lambda > 0$ indicates that expected utility would decrease as variance increased, representing a risk-averse individual. Appendix A discusses the relationship between λ and Arrow-Pratt measures of absolute and relative risk aversion in more detail.

The next section develops demand curves for contracts for the three categories of consumers described in Section 9.3. Implicitly, we assume that consumers use the variance of profit as their measure of risk, and that their optimal hedging decisions, given their attitude to risk, can be found using a mean-variance vN-M function. It will be assumed that the various factors driving the degree to which firms are concerned about spot price risk (e.g., relative significance of electricity costs, and risk attitude of the decision makers) can be combined into a single risk-aversion parameter, as described above.

9.5.2 Terminology

Our analysis considers the following general scenario. Firms' electricity requirements, within the given LDC subperiod (when applicable), are denoted by l, which may be uncertain within the subperiod ($\tilde{l}$). Firms either purchase their requirements from the spot market at an uncertain price $\tilde{p}$, a random variable, or have the option to purchase k units on a CfD, which has strike price f. Consumers with unresponsive loads will be indexed by j, while those with responsive loads will carry index i.

9.5.3 Consumers with Unresponsive Loads

First, we will reproduce the standard theory ((see Guthrie (1998) for a full treatment) for a firm whose electricity load is a random variable. As discussed above, representing load with a random variable $\tilde{l}$ implies that at least some of the variability in the firm's electricity requirements is unpredictable, and presents a risk to the firm. Later, we will assume that l varies deterministically, implying that load is fixed and certain within the LDC subperiod.

Given that the firm does not respond in a profit maximising fashion to movements in the electricity price, we can ignore firm profits and form the following function for (uncertain) net electricity cost, $\tilde{C}_j$ for consumer j:

$$\tilde{C}_j = (\tilde{l}_j - k_j)\tilde{p} + k_j f \tag{9.4}$$

Thus, the expected value and variance of total electricity cost are, respectively:

$$E\left[\tilde{C}_j\right] = E\left[\tilde{l}_j\tilde{p}\right] + k_j f - k_j E\left[\tilde{p}\right] \tag{9.5}$$

$$Var\left[\tilde{C}_j\right] = Var\left[\tilde{l}_j\tilde{p}\right] + k_j^2 Var\left[\tilde{p}\right] - 2k_j Cov\left[\tilde{l}_j\tilde{p}, \tilde{p}\right] \tag{9.6}$$

Using equation (9.3), we can form the mean-variance optimisation problem for the firm. Since the traditional form of the mean-variance objective assumes that the firm is maximising the tradeoff between expected wealth and variance of wealth, we let $E\left[\tilde{x}\right] = -E\left[\tilde{C}\right]$, i.e., we ignore any benefit the unresponsive consumers may derive from their electricity purchases. Dropping, for now, the index j, firms maximise the following objective function:

$$E\left[U\left(\tilde{C}\right)\right] = -E\left[\tilde{l}\tilde{p}\right] - kf + kE\left[\tilde{p}\right] - \frac{\lambda}{2}\left(Var\left[\tilde{l}\tilde{p}\right] + k^2 Var\left[\tilde{p}\right] - 2kCov\left[\tilde{l}\tilde{p}, \tilde{p}\right]\right) \tag{9.7}$$

The appropriate first order condition for the firm is:

$$0 = \frac{dE\left[U\left(\tilde{C}\right)\right]}{dk} = E\left[\tilde{p}\right] - f - \lambda k Var\left[\tilde{p}\right] + \lambda Cov\left[\tilde{l}\tilde{p}, \tilde{p}\right] \tag{9.8}$$

Thus the consumer will maximise expected utility by purchasing

$$k = \frac{E\left[\tilde{p}\right] - f}{\lambda Var\left[\tilde{p}\right]} + \frac{Cov\left[\tilde{l}\tilde{p}, \tilde{p}\right]}{Var\left[\tilde{p}\right]} \tag{9.9}$$

units of electricity on contracts for differences. This shows that demand for contracts is made up of two components, reflecting the profit maximising and variance minimising incentives of mean-variance preferences. The first term in (9.9) is commonly referred to as "speculative hedging", reflecting the profit maximising (or, in this case, cost minimising) objective. If the firm believes the spot price will be lower than the contract price, it will purchase fewer contracts in order to take advantage of this. On the other hand, if spot prices are expected to rise above the strike price, these consumers will become over-contracted, planning to sell back electricity to the contract market at the favourable expected price. This speculative behaviour is moderated by the firm's risk aversion, and the volatility in the spot price.

The second component indicates that the consumer is also hedging the risk in the expected total cost of their load, $E\left[\tilde{l}\tilde{p}\right]$. If the total spot cost is positively correlated with the spot price, i.e., when the spot price is high, the spot value of their demand is high, they will purchase more contracts in order to hedge this. When the spot price is low, relative to the contract price, the consumer will make an effective loss on the contracted units, but can most afford to do it in these states. Equivalently, if the spot price is high, the firm will make an effective gain on the contracted units, but this will come at a time when it most needs it, since the total cost of its load is high (Guthrie (1998))[61].

Rearranging (9.9) we get:

$$f = E\left[\tilde{p}\right] + \lambda\left(Cov\left[\tilde{l}\tilde{p}, \tilde{p}\right] - kVar\left[\tilde{p}\right]\right) \tag{9.10}$$

If electricity load, within the given subperiod, is known, this reduces to:

$$f = E\left[\tilde{p}\right] + \lambda Var\left[\tilde{p}\right](l - k) \tag{9.11}$$

or, rearranging for contracts

[61] The effect of load uncertainty on its own is unclear, since the covariance of any two random variables is difficult to interpret in terms of the variance of the underlying variables. Hull (1995) describes one of the major disadvantages of covariance is the inability to read anything into its value, or interpret the likely effect of any change in the input distributions.

$$k = l + \frac{E[\tilde{p}] - f}{\lambda Var[\tilde{p}]} \tag{9.12}$$

Equation (9.12), the standard portfolio hedging equation, indicates that consumers who wish to maximise the expected M-V function of their electricity costs, with known load, will purchase contracts equal to their demand for electricity, adjusted by a speculative element which is moderated by their aversion to the spot variance. The speculative element is identical to, and thus has the same interpretation as that for an uncertain load.

Importantly, the unresponsive firm's demand curve for contracts is downward sloping in contracts, with slope $-\lambda Var[\tilde{p}]$. Thus the risk premium (the excess over the expected spot price that consumers are willing to pay to avoid risk) decreases as contracts increase, i.e., the amount of unhedged load decreases. If the firm purchases contracts to match its load, it will not be willing to pay any more than the expected spot price for additional contracts.

The result (9.12) differs slightly from that of Powell (1993), who assumed that electricity consumers were large enough to act strategically. These firms were aware that an increase in the aggregate level of contracting would drive future spot prices down. Hence the denominator in the second term of (9.12) included a partial derivative reflecting this. We assume that consumers are too small to individually have this effect on price.

9.5.4 Consumers with Responsive Loads

As discussed in Section 9.3, a firm may possess control over its electricity load by altering its level of activity. Here we consider firms that can change their activity level in response to unpredictable changes in the unit cost of electricity, i.e., the spot price.

Firstly, we must develop expressions representing the optimal quantity response of the firm to a change in the cost of electricity. These equations can then be used to examine how a risk-averse firm behaves in a contract market.

Marginal Benefit of load

In order to build a model of how the individual firm will adjust its electricity load in response to a change in the electricity cost, we make the following assumptions:

1. We can represent the marginal benefit or valuation (v_i) of a unit of load (l_i) to firm *i* with a function $v_i = f(l_i)$.
2. The above function is the result of the firm acting optimally in its output market, so the integral of $f(l_i)$ is the total benefit to firm *i* for demanding l_i units of electricity.
3. $f(l_i)$ is known, linear, and downward sloping, with respect to l_i.

(3) is true if we assume that electricity load is related positively and linearly to the activity level of the firm (which seems reasonable if electricity is a variable cost of production, and total benefit is quadratic in the level of activity). Hence by the optimality conditions on activity level, implied by (2), the marginal benefit of load is decreasing in l_i. Everything we need to know about the output market that an individual firm faces is included in $f(l_i)$. Let:

$$v_i = f\left(l_i\right) = a_{0i} - \rho_i l_i \tag{9.13}$$

where a_{0i} and ρ_i are constants.

Note that $f(l_i)$ effectively represents the firm's individual demand curve for electricity. Hence v_i is the price it would be willing to pay to procure l_i units of electricity, ignoring the effects of risk aversion. Rearranging (9.13), we can say that for a given electricity price, *p*, the firm's demand for electricity is:

$$l_i = \frac{a_{0i} - p}{\rho} \tag{9.14}$$

The net profit to the firm for a given activity level is the area under $f(l_i)$, between 0 and the value of l_i implied by that activity level, less the total cost of electricity purchases. In the case where total electricity load is purchased at price p, the total cost of electricity purchases is $l_i p$, or area B in Figure 9.5. The net profit corresponds to area A.

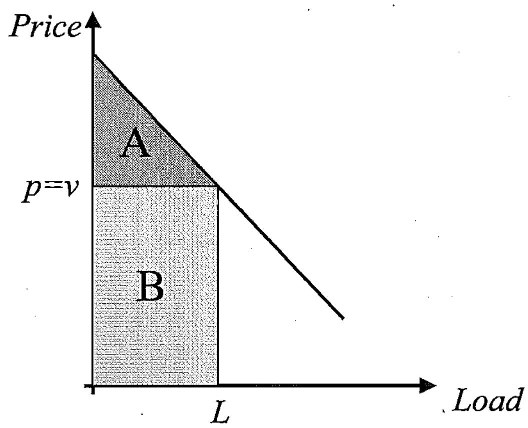


Figure 9.5 Total benefit, total cost and net profit from purchasing electricity

Hence we can describe the area in the triangle depicted in Figure 9.5 above as

$$\begin{aligned}\pi_i(p) &= \frac{(a_{0i} - p)l_i}{2} \\ &= \frac{(a_{0i} - p)^2}{2\rho_i}\end{aligned} \tag{9.15}$$

Profit maximisation with contracts

The firm has two available sources of electricity: to sign a certain quantity k_i at strike price f on contract, and/or purchase off the uncertain spot market. The problem for the firm is to determine how its profit maximising activity level responds to a varying spot price, when a certain quantity of its load may be purchased on contract. This, in turn, will help us determine the firm's demand for contracts at a given contract price, using a mean-variance function.

Consider the situation where a firm has purchased a quantity of contracts, k_i. In a given period t, the firm observes a spot price, p^t. The firm might choose to increase production if p^t was low, or if p^t was high, production, and thus load, could be decreased. If its electricity load was reduced below its contract level, it would effectively sell back electricity to the spot market, at a potential profit (if the spot price exceeded the contract price). Figure 9.6 & Figure 9.7 illustrate these situations, respectively, and we will now deal with each of them in detail.

For a low spot price in period t, p_t^L, the firm might consider increasing its production and thus load to that implied by the demand curve for electricity, l_i^t. The contract energy was paid for at the contract price f. Note that f will include the effects of risk aversion, so will not necessarily correspond to $v(k_i)$, the marginal benefit of producing where $l_i = k_i$. The firm accrues a total benefit, from the contract units, equal to the area under $f(l_i)$ between 0 and k_i.

The firm will purchase the net difference between the contract quantity and the total load, $l_i^t - k_i$, off the spot market. This additional energy is paid for at the spot price, and these units accrue a net profit to the firm of area B.

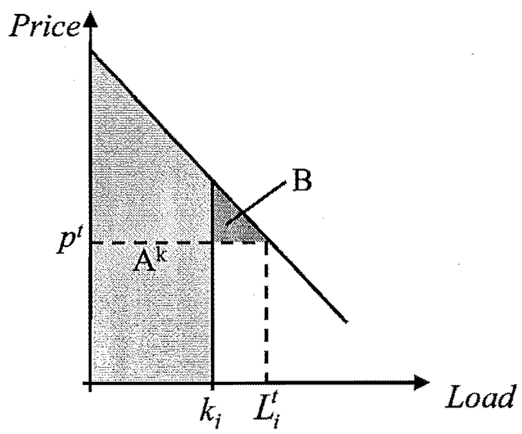


Figure 9.6: Load for low spot price

Figure 3 illustrates the case where the firm observes a high spot price, p_t^H that would imply an optimal activity level corresponding to an electricity demand less than the contract quantity. Through the difference payments made under the CfD format, the firm consumes l_i units of electricity, and effectively sells back (k_i - L_i) to the spot market at the prevailing price p, making a net contract profit of (p - f)(k_i - l_i) in that period.

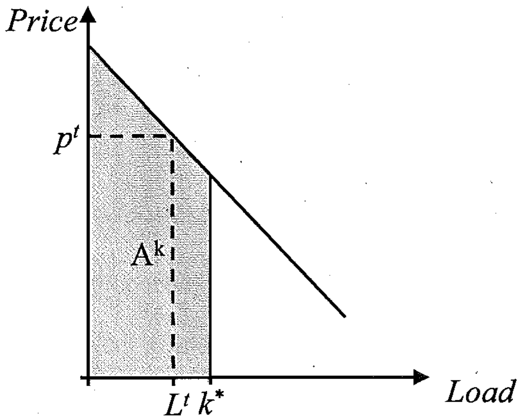


Figure 9.7 Load for high spot price

In either case, the firm realises a total benefit from operating at a level corresponding to a load of l_i^t units of electricity of:

$$B_i^t = l_i^t p^t + \frac{l_i^t \left(a_{0i} - p^t\right)}{2} \tag{9.16}$$

Knowing the relationship between optimal load response and price, we can substitute (9.14) into (9.16):

$$B_i^t = \left(\frac{a_{0i} - p^t}{\rho}\right) p^t + \frac{\left(a_{0i} - p^t\right)^2}{2\rho} \tag{9.17}$$

The total contract and spot cost to the firm requiring l_i^t is:

$$C_i^t = k_i f + \left(\frac{a_{0i} - p^t}{\rho} - k_i\right) p^t \tag{9.18}$$

Net benefit is thus

$$\begin{aligned}\Pi_i^t &= B_i^t - C_i^t \\ &= \left(\frac{a_{0i} - p^t}{\rho}\right)p^t + \frac{\left(a_{0i} - p^t\right)^2}{2\rho} - k_i f - \left(\frac{a_{0i} - p^t}{\rho} - k_i\right)p^t \\ &= \frac{\left(a_{0i} - p^t\right)^2}{2\rho} - k_i\left(f - p^t\right)\end{aligned} \tag{9.19}$$

Comparing 9.19 with 9.15, where contracts were not considered, it is interesting to observe that 9.19 implies that profit maximising load (and thus output) decisions are entirely separable from profit maximising contract decisions[62], in the situation where a firm intends to take full advantage of the profit making opportunities implied by the CfD. Such a separation would be evident in a management structure where a manager of plant simply responds to spot variations by setting the output to maximise spot profits, while a higher-level contract manager determines the contract level, and no communication is made from high to low level. Equation 9.19 shows that the result of this management structure would be the same as when contracts and load are jointly optimised. This is intuitive, as a feature of the financial contract is that the signals for load management provided by spot prices are preserved. Hence whether a firm alters its production in response to spot prices in order to remain in "output optimality" when not hedged, or to make a contract profit when they are hedged, the manner in which production is changed is exactly the same.

Uncertainty and risk aversion

In order to represent risk-averse behaviour, we substitute 9.19 into the mean-variance equation described in Section 9.5:

$$U(\pi) = E\left[\frac{\left(a_0 - \tilde{p}\right)^2}{2\rho} - kf + k\tilde{p}\right] - \frac{\lambda}{2}Var\left[\frac{\left(a_0 - \tilde{p}\right)^2}{2\rho} - kf + k\tilde{p}\right] \tag{9.20}$$

62 Although, it is worth noting that this assumes that the output decisions of the firm have no effect on the spot price.

$$= \frac{E\left[\left(a_0 - \tilde{p}\right)^2\right]}{2\rho} - kf + kE\left[\tilde{p}\right] - \frac{\lambda}{2}k^2 Var\left(\tilde{p}\right) - \frac{\lambda}{8\rho^2}\left(4a_0^2 Var\left(\tilde{p}\right) + Var\left(\tilde{p}^2\right)\right)$$
$$+\frac{\lambda}{2}\left[\frac{a_0 k}{\rho}2Cov\left(\tilde{p},\tilde{p}\right) + Cov\left(\tilde{p},\tilde{p}^2\right)\left(\frac{a_0}{\rho} - \frac{k}{2\rho}\right)\right]$$

Differentiating with respect to contracts and setting to zero:

$$0 = E\left[\tilde{p}\right] - f - \lambda k Var\left(\tilde{p}\right) + \frac{\lambda}{2}\left[\frac{a_0}{\rho}2Var\left(\tilde{p}\right) - Cov\left(\tilde{p},\tilde{p}^2\right)\left(\frac{1}{2\rho}\right)\right] \tag{9.21}$$

Rearranging for contracts:

$$k = \frac{E\left[\tilde{p}\right] - f}{\lambda Var\left[\tilde{p}\right]} + \frac{a_0}{\rho} - \frac{Cov\left[\tilde{p},\tilde{p}^2\right]}{4\rho Var\left[\tilde{p}\right]}$$

or

$$f\left(k\right) = E\left[\tilde{p}\right] + \lambda Var(\tilde{p})\left(\frac{a_0}{\rho} - k\right) - \frac{\lambda}{4\rho}Cov\left(\tilde{p},\tilde{p}^2\right) \tag{9.22}$$

Equation (9.22) shows that, in addition to the expected spot price, firms are willing to pay a risk premium determined by both their exposure to the variance of the spot price, and a covariance term. The second term represents the maximum possible exposure they might face (their profit-maximising load under a zero-spot price scenario) for that particular contract level, in the same way that the unresponsive consumers' risk premium was determined by their residual exposure to the spot price, for a given level of contract cover.

However, the responsive consumer's level of contracts is also determined by the covariance of the spot price with its square. This term originated in the objective function (9.20) as the co-variability of the spot price with total profit (via the profit maximising load response). Total profit for the firm was a function of the electricity spot price, or more exactly, a decreasing function of $\tilde{p}^2$, when the firm responded optimally to the spot price. Hence higher spot prices lead to higher values of $\tilde{p}^2$, and lower optimal profits. Thus by responding to the electricity spot price, firms are performing their own method of profit stabilisation: high spot prices lead to low profits, and low prices lead to high profits. Since the CfD structure leads firms to respond to varying spot prices in the same way regardless of the contract level (see above), this hedging is unaffected by the particular level of contracts purchased by the firm.

The degree to which this behaviour stabilises profits is a function of both the effect of $\tilde{p}^2$ on profits, and the relationship between $\tilde{p}$ and $\tilde{p}^2$. This is why we see both the marginal change in load, ρ (relating load to price and thus profits), and the covariance of price with its square, in the optimal hedging equation. As expected, since the covariance term is positive in almost all reasonable price distributions[63], this leads to a lower level of optimal contracting than for their unresponsive counterparts. In fact, for levels of contracts that cause the third term to exceed the second term in (9.22), generators may have to offer contracts at a discount to the spot price in order to induce responsive consumers into a higher level of contracting.

9.5.5 Industry Demand for Contracts

Now that expressions for the optimal level of hedging for each type of individual consumer have been developed, we can aggregate them to form an industry DCC.

Total market demand for contracts is defined by:

$$K = \sum_j k_{U,j} + \sum_i k_{R,i} \tag{9.23}$$

where $k_{U,j}$ is the contract demand of the *j*th consumer with an unresponsive load, and $k_{R,i}$ is the contract demand of the *i*th consumer with a responsive load.

[63] The only situations in which this term would be negative is a distribution in which a significant proportion of prices were less than 0.

Initially, we will sum each type of demand separately. We assume that all load variability of the unresponsive consumers is predictable, hence equation (9.12) is appropriate, and l is therefore the average load within the given LDC subperiod. Summing across all unresponsive consumers:

$$\begin{aligned}\sum_j k_{U,j} &= \sum_j \left(\frac{E[\tilde{p}]-f}{\lambda_j Var[\tilde{p}]} + l_{U,j} \right) \\ &= \frac{E[\tilde{p}]-f}{Var[\tilde{p}]} \sum_j \frac{1}{\lambda_j} + L_U \end{aligned} \tag{9.24}$$

where $L_U = \sum_j \tilde{l}_{U,j}$.

Assuming each responsive consumer purchases contracts according to (9.21), total contract demand for this consumer type is:

$$\sum_i k_{R,i} = \frac{E[\tilde{p}]-f}{Var[\tilde{p}]} \left(\sum_i \frac{1}{\lambda_i} \right) + \sum_i \frac{a_{0,i}}{\rho_i} - \frac{Cov\left[\tilde{p},\tilde{p}^2\right]}{4Var[\tilde{p}]} \sum_i \frac{1}{\rho_i} \tag{9.25}$$

If we assume that all responsive consumers are identical, we can simplify (9.25) further. The aggregate spot demand from responsive consumers can be found by rearranging (9.14), and summing across all such consumers:

$$\begin{aligned}\sum_i l_{R,i} &= \sum_i \frac{a_{0,i}}{\rho_i} - \tilde{p} \sum_i \frac{1}{\rho_i} \\ &= \frac{n a_{0,i}}{\rho_i} - \tilde{p} \frac{n}{\rho_i}\end{aligned}$$

Let $\rho_i / n = b$ and $a_{0,i} = A$. Equation (9.25) can now be written

$$\sum_i k_{R,i} = \frac{E[\tilde{p}]-f}{Var[\tilde{p}]} \left(\sum_i \frac{1}{\lambda_i} \right) + \frac{A}{b} - \frac{Cov\left[\tilde{p},\tilde{p}^2\right]}{4bVar[\tilde{p}]} \tag{9.26}$$

(As an aside, total spot demand, $\tilde{D}$, is obtained by adding in the unresponsive loads:

$$\tilde{D} = \frac{na_{0,i}}{\rho_i} - \tilde{p}\frac{n}{\rho_i} + L_U \tag{9.27}$$

The aggregate inverse spot demand curve is thus defined by

$$p(D, L_U) = A - b(D - L_U) \tag{9.28}$$

(9.28) is consistent with the inverse spot demand curve that will be used by generators in their profit maximisation (Chapters 10 and 11)).

We can now sum across both types of consumer, using (9.23), (9.24) and (9.26):

$$\begin{aligned} K &= L_U + \frac{E[\tilde{p}] - f}{Var[\tilde{p}]}\left(\sum_{i,j}\left(\frac{1}{\lambda_i} + \frac{1}{\lambda_j}\right)\right) + \frac{A}{b} - \frac{Cov[\tilde{p}, \tilde{p}^2]}{4bVar[\tilde{p}]} \\ &= \frac{A}{b} + L_U + \frac{E[\tilde{p}] - f}{\Lambda Var[\tilde{p}]} - \frac{Cov[\tilde{p}, \tilde{p}^2]}{4bVar[\tilde{p}]} \end{aligned} \tag{9.29}$$

where $\frac{1}{\Lambda} = \sum_{i,j}\left(\frac{1}{\lambda_i} + \frac{1}{\lambda_j}\right)$, i.e., the harmonic average of all consumers' risk aversion.

Rearranging to make *f*, the contract price, the subject, our total DCC is:

$$\begin{aligned} f(K) &= E[\tilde{p}] + \Lambda Var[\tilde{p}]\left(\frac{A}{b} + L_U - K\right) - \frac{\Lambda}{4b}Cov[\tilde{p}, \tilde{p}^2] \\ &= E[\tilde{p}] - \Lambda Var[\tilde{p}]K + \Lambda\left(Var[\tilde{p}]\left(\frac{A}{b} + L_U\right) - \frac{1}{4b}Cov[\tilde{p}, \tilde{p}^2]\right) \end{aligned} \tag{9.30}$$

9.6 Conclusions

This chapter presented models of the various types of demand for CfD's, by consumers of the electricity commodity. Demand functions were constructed by considering how consumers' load responded to a change in the cost of electricity, if at all. In recognition of the wide variety of electricity purchasing behaviour that exists in the market, consumers were categorised into two groups: those that had loads which did not alter in response to the unit cost of electricity, and those who could perform some degree of load control as the electricity price changed, and did so in a profit maximising manner. No attempt was made to optimise the actual response process; it was assumed that all information pertaining to the profit maximising behaviour of consumption was modelled by their individual demand curves for the electricity commodity.

It was assumed that all consumers in the contract market were risk averse, and utility maximisation was used to analyse consumers' contract buying behaviour in the face of volatility in the electricity spot price. This volatility was measured by the variance of the spot price within a PDC subperiod, the latter describing all the known information about the overall variability of the price. Also it was assumed that no individual consumer anticipated the effect their decision had on total contract demand.

Their decision making under risk was represented by a mean-variance maximisation, and the degree of aversion to the variance of their cashflows could be captured by a single risk aversion parameter.

The result of the M-V maximisation was two general equations representing the demand for contracts, for each type of consumer, as a function of contract price, the mean and variance of the spot price, and other parameters pertaining to the behaviour of their electricity demands. Both demand curves shared a similar property: the price consumers were willing to pay for a given quantity of contracts was the expected spot price, plus a risk premium.

The driving factors of the risk premium distinguished the two types of consumer. For the cost minimising consumer, whose load varied in a predictable fashion, this premium was determined by the product of the variance of the spot price, and that portion of their predicted load that was still exposed to the spot market for a given contract quantity. This product was then weighted by their risk aversion. (If load was not known with certainty, a more complicated expression for this premium was developed, by adjusting the net exposure to variance for the fact that load may in fact co-vary with the spot price, which could either increase or decrease the net risk faced by the consumer for a given commitment to a contract level. However, we have chosen not to pursue this case any further).

A profit maximising firm which had the option of varying its activity level, and thus load, in response to the spot price, formed risk premiums differently. With the responsiveness of their loads, these consumers effectively had a second risk-hedging tool at their disposal – the effect of a change in electricity costs could be partially mitigated by increasing or reducing their electricity load. The risk premium included the spot price hedging effect of contracts in a similar way to the cost minimising consumer above, but the nature of the load response was represented by an extra term that reduced the premium. This term reflected the fact that a firm would respond to high (low) prices with low (high) optimal load, and that this would stabilise profits. Hence we have the intuitively reasonable conclusion that consumers who exhibit optimally-responsive electricity demands are less willing to pay to avoid spot price variance than those who exhibit fixed loads, and thus can't respond.

In reality, there would be a range of firm behaviour in a single electricity market. Many firms may have purchased electricity supply contracts that allow a variable amount of electricity at a fixed spot price, up to a certain quantity, and over that quantity a penalty or additional price is imposed (swing options). Other firms may exhibit similar behaviour, but may alter production in response to the wholesale market when spot prices reach extreme values, in particular, in times of electricity shortages when the conditions in the electricity supply contract require firms to reduce load, or firms are encouraged to by other market participants.

However, as has been previously mentioned, it is not the purpose of this chapter, nor this thesis, to provide an exhaustive treatment of the demand side of the electricity market, for either the commodity or contracts. This chapter intends to produce a continuous demand curve for contracts, which enables a generation firm to account for the likely response of hedging consumers to a change in the quantity or price of contracts it makes available.

It appears from equations (9.12) and (9.21) that estimating the price-quantity relationship in the spot market under the assumptions employed here may not be difficult. Assuming the generator already has a relatively accurate idea of the spot demand curve, there is little additional knowledge of demand-side behaviour required. By making full use of the information about firms' likely response to changing spot prices in the demand curve, we were able to construct expressions for contract demand that are closely tied to spot market behaviour, thus reinforcing the contract-spot relationship postulated in the framework for this thesis detailed in Chapter 3. Since the demand for contracts actually reduces to a number of terms involving the spot price, dominant generation firms may actually have a degree of control over the position of the contract demand curve, as well as the position they select on the contract demand curve. This issue is pursued further in Chapter 11. Before that, Chapter 10 provides a mathematical model of both contract and spot markets simultaneously, incorporating the strategic behaviour of dominant firms and the natural variability in the market.

10

A MODEL OF JOINT SPOT-CONTRACT EQUILIBRIA

10.1 Introduction

Previous chapters have established principles essential for modelling firm behaviour according to the risk management triangle presented in Chapter 3. The framework proposed by the triangle suggested three inter-related mechanisms were available to a large hydro firm to manage risk: market power, long-term contracts, and storage.

Chapter 8 showed that, for a particular system, market power and storage can be effectively used by Cournot firms to mitigate the effects of hydrological uncertainty on firm profits. Here, for a firm that had sold few long term contracts, risk (the standard deviation of profits) was reduced by up to 60% (when compared with a competitive firm) by the use of a combination of market power and storage. This risk "control" was not achieved at the expense of profit, either, as firms were assumed to be risk neutral.

The analysis also showed that, for the hydro firm, profit risk was increasing in the level of contracts, since forward arrangements reduce the incentive to use market power, and thus increased profit risk for a dominant firm. This is contrary to the conventional wisdom that long term contracts reduce risk.

If dominant firms experience satisfactory levels of risk by acting in a risk-neutral, profit maximising manner, it seems reasonable to simplify the analysis by continuing to model the risk attitude of these firms in this way. It is reasonable to expect that such firms will only sell forward contracts if it is profitable to do so. However, the model used in Chapter 8 assumed that contracts were valued at the mean spot price, implying that the electricity supply firms received no premium from the consumers for output sold to them on long-term CfD's. As suggested earlier, long-term forward contracts may, in fact, attract a premium if consumers are sufficiently risk averse. This is particularly true of countries where contract markets are relatively under-developed, and thus contracts are not completely standardised or freely traded by risk neutral speculators.

Hence, in Chapter 9, a possible model of the behaviour of electricity consumers in a contract market was developed, in order to find a reasonable estimate of the premiums risk averse consumers would be willing to pay. Here the term 'market' loosely refers to the process by which producers and consumers of electricity agree on the terms of a financial contract. It was shown that risk averse consumers of electricity find forward contracts attractive, to hedge risk, but the magnitude of the risk premium they are willing to pay, to avoid the variability of the spot price, depends on their own ability to "naturally hedge" the variations, by adjusting their own electricity requirements in response to the spot price. Hence the greatest risk premiums were offered by those consumers who either chose not to, or could not, adjust their loads in response to the spot price, while those who could adjust their demand within a reasonable timeframe offered a lower premium to be committed to a long term arrangement.

If, as suggested in Chapter 8, the uncertainty introduced through inflow variations is adequately managed via the natural decision making process of a generation firm with market power, the risk management triangle reduces to determining the best use of contracts and market power, assuming the storage resources are managed in a similar fashion to the reservoir management policies developed by Scott (1998). The direct interactions of these tools, under risk neutrality, have been outlined in earlier chapters and could be summarised as follows:

- Profit maximising generation levels are increasing, and thus spot prices decreasing, in the quantity of contracts sold by the firm, when no feedback from the spot market outcomes, on contract prices, is accounted for.

- Profit maximising contract levels, in turn are determined (amongst other things) by the consumers' risk aversion, expectations of spot price, and observed variance of the spot price.

These interactions make it clear that optimal levels of contracts and generation must be determined simultaneously. As pointed out in Chapter 7, while a number of authors have, across a variety of applications, considered the effect of contracts on optimal spot market behaviour, few have considered the feedback loop from the spot market to the formation of expectations used by consumers in the contract negotiation process. In many situations this is not relevant, as contract prices are set beyond the control of individual firms. However, this thesis considers those cases where supply firms have a significant amount of influence over contract equilibria, and thus emphasises the effect that spot market outcomes have on the demand for CfD's.

Initially, however, this chapter will begin this process by developing a model where generators ignore, in their *spot* market optimality conditions, the effect that their generation behaviour has on the contract demand curve. That extension is made in Chapter 11. This chapter presents a model where firms determine output levels, for a range of cost "states", in a short-run profit maximising manner, and without deliberately trying to influence the contract market. The firm must also make a single, expected-profit maximising contract decision, given that any of the cost states may be observed. Since we are interested in the equilibrium between these two markets, these decisions are connected by the distribution of spot prices resulting from the spot market equilibria, which consumers used in the development of their demand curve for contracts. This process is illustrated in Figure 10.1.

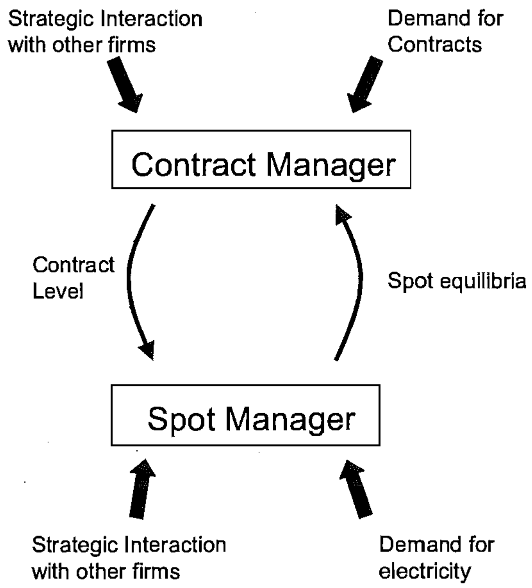


Figure 10.1 "Naïve" spot-contract decision-making

It is clear that, even though this model is relatively simplistic, it includes significant advances on the work of Green (1993), Allaz (1992) and Powell (1993), who also examined the joint spot-contract decision making problem under Cournot assumptions. First, we consider how a range of generator cost states, and the generators' profit maximising response to it, creates a spot price distribution. Second, we wish to find the equilibrium between the two markets, so we will simultaneously solve the spot and contract market optimality conditions, which, in the chapter, are connected by the spot price distribution. Third, and perhaps most importantly, the next chapter will consider how generators may manipulate this connection between the markets to their own advantage.

The model of the spot market and the generators' optimality conditions are developed in Section 10.4, while a similar analysis for the contract market is presented in Section 10.5. The optimality expressions are then combined, in Section 10.6, to form the multi-state, joint-equilibria model.

Firstly, however, we will introduce the concept of long-run equilibrium, which will underlie the models that are presented in the remainder of this thesis.

10.2 Inter-temporal Concepts

As discussed in earlier chapters, much of the motivation for deregulation comes from the desire to reduce the market power of large supply firms, or at least to remove the incentives to abuse that power. If firms are able to exert market power, it is reasonable to believe that prices are above marginal cost and thus the electricity dispatch is less efficient than it could be. It is expected that by providing the correct price signals through market mechanisms, prices and output decisions will move closer to the perfectly competitive equilibrium, either by incumbents increasing output, or by the entry of cheaper plant into the market. While the reality is that, in a corporate environment, we will never have perfect information about companies' true underlying costs, a "pseudo-competitive" equilibrium could be loosely defined as where the price consumers pay for electricity is close to the best available knowledge of the marginal cost of electricity production[64].

This suggests that there are two critical elements to an analysis under this framework - the equilibrium market state in the long run, and the process by which the market makes the transition from whatever state it is in at present to the long-run state. We will now present a brief discussion of each aspect. As shown below, the transitory "dis-equilibrium" phase is outside the scope of this thesis, and will not be modelled here, but the long-run equilibrium state provides fertile ground for modelling the problem we have represented with the risk management triangle.

10.2.1 Short-run Dis-equilibrium

A dis-equilibrium analysis would attempt to analyse how dominant firms interact with potential entrants, consumers, and rivals in both the contract and spot market, under a number of uncertainties. While most established firms will have already obtained a significant amount of knowledge of input issues, e.g., hydrology and cost of fossil fuels, during this period of "dis-equilibrium" most market participants are gaining new information about the response of the **market** to different courses of action, and how it influences established characteristics of the industry (e.g., transmission reliability and load growth). The emphasis during this phase is on how firms adjust their behaviour over time as this new information is received. Supply firms increase their knowledge of the reaction of their rivals and consumers to their decisions. Consumers and retail firms learn of spot price behaviour in response to different seasons, economic and regulatory environments. As an example, Helm and Powell (1992), and Lowrey (1997) both provide an empirical analysis of how dominant generators' spot market behaviour changed over time in response to different regulatory actions and the expiry of vesting contracts in the England and Wales electricity market.

Of critical importance to all parties during the dis-equilibrium stage is how the behaviour of the market is affected by, and affects, contract negotiations. Firms are not only learning how to manage their own risk in the new environment, but also how other firms respond to risk. In particular, firms selling contracts will try to determine the extent to which consumers are willing to pay to hedge risk. While the measurement of risk aversion is an inexact science (see Chapter 2 and Appendix A for a fuller discussion), supply firms may experiment over time, through the contract negotiation process, in order to obtain estimates of how consumers measure and respond to risk.

As suggested above, the move towards a more efficient market may include the encouragement of new, efficient firms to enter the industry. Faced with this prospect, incumbent firms would evaluate the likely state of the market in the long run, with and without potential entrants, to determine the profitability of each scenario. From this, they would decide whether they should allow entry to naturally occur, or if some form of entry deterrence is desirable. The most obvious entry-deterring strategy is for incumbent firms to ensure the average spot price does not exceed a limit price (a price that would trigger entry; usually either the average or marginal cost of the entrant). More generally, the incumbent would ensure that the profits accruing to a potential entrant were not sufficient to recover the fixed (sunk) costs of entry. For example, the potential entry of a high marginal cost thermal may be deterred by minimising the hours in which it would be

[64] The issue of whether this refers to short-run or long-run marginal cost was briefly addressed in Chapter 4.

called on to generate, which could be achieved by an incumbent increasing output in high-demand periods. Most entry-deterring strategies are usually sub-optimal from a short-run profit maximising (SRPM) point of view[65], but are desirable from the firm's perspective when the alternative (i.e., allowing the firm to enter the market) is less profitable in the long run. Newbery (1998) points out that if contracts can be issued to finance entry for a generator, it is no longer the spot price that must be limited to deter entry, but the contract price as well. However, other authors (see Aghion and Bolton (1987)) contend that contracts can equally deter entry, if potential consumers for the entrant are already committed to any of the incumbents through long-term contracts, or supply arrangements with vertically integrated retail companies. Unless the consumer can easily trade such contracts, the potential market for a potential entrant is drastically reduced. Powell (1993) also argued that if the industry has a high level of contracts in effect, the entrants' assessment of its profitability is potentially compromised, since the spot price does not accurately convey the true price being paid for electricity. If these contracts aren't transparent to all participants, the potential exists for incumbents to use the spot price to deter entry, without sacrificing their own profits.

It is clear that the dis-equilibrium state of the market involves the combination of a number of complex issues which are outside the scope of this thesis. Conjectural variations other than the Cournot model already chosen, and the vast area of competition under imperfect information are aspects worthy of investigation in future development of the models presented in this thesis, but will not be discussed further here. Instead, we will concentrate on a market characterised by a long run equilibrium.

10.2.2 Long-run Equilibrium

In order to avoid the complexities of dis-equilibrium analysis, we can instead base our modelling on a convenient, hypothetical state of long run equilibrium, which could be interpreted as the best estimate of the market state in the future, rather than a prediction of

[65] Except in the case where the SRPM solution results in a price too low for the potential entrant. In such a case, entry is usually said to be blockaded (Bain, 1949)

what will happen in any particular year. The concept of equilibrium that we will use is characterised by the following:

- A minimum of political or regulatory intervention, with little chance of change in the near future,
- Stability in technological, social, and economic factors, and
- No imminent exhaustion of resources (e.g., fossil fuels) or fatigue of generation plant.

The equilibrium may or may not be the desired competitive one, but is mainly typified by stable spot and contract market behaviour, by all participants. This allows us to assume that there is "full information" in the market, an assumption that is critical to the models that follow. This assumption has two important implications. First, consumers have a reliable, accurate sample of generator behaviour that represents all the possible market outcomes that can occur. While uncertainty may still exist as to which outcome will be realised in a given year, the consumers know that it must be selected from the distribution of outcomes they have already observed. This avoids the "surprise" factor of the dis-equilibrium state of the market.

Second, generators are aware of the factors that drive consumers' spot and contract behaviour, i.e., they have good estimates of the nature of responsive, and unresponsive load, and the incentives acting on consumers to purchase forward contracts. Thus, they know, or have reasonably accurate estimates of, the form of the DCC, developed in Chapter 9.

These assumptions allow all market participants to make full use of the knowledge they have obtained of recent market behaviour, in order to make predictions about possible future states of the market. That is, the market has "settled down" to a degree where this information is a good predictor of the near future. In order to obtain a tractable model that supports the assumption of a general equilibrium and full information, two further issues must be ignored, namely load growth and the prospect of entry. Load is assumed not to grow over time, mainly for computational convenience. Its inclusion in the model would be possible, but would potentially introduce another dimension of uncertainty, and require the consumers to update their estimates of the expected spot price behaviour over time.

It is important to emphasise at this point that, as discussed above, the equilibrium conditions derived will not consider any explicit inter-temporal effects. Given the definition of long-run equilibrium presented above, all parameters involved in the participants' decisions are stable with respect to time, and in this sense the model is "timeless", or, at least, describes the market equilibrium that would be observed in a general time period. Hence, in the discussion and analysis that follows, the term "period" will refer to a particular "round" of the contract and spot decisions described by the equilibrium. In every period (even though the equations will only examine a general period), consumers and generators alike face an uncertain distribution of "states", select an equilibrium contract level, and then the uncertainty will be resolved, allowing participants to make profit maximising spot decisions. Of course, once these decisions are made and acted upon, the model does, implicitly, move to the next time "period" and begin the same process. However, since the distribution of *states* that might be observed in any period remains the same, regardless of which time period the contract decision is made in, the exact same equilibrium decisions will be observed in that period, and indeed every subsequent period, reflecting the long-run equilibrium concept described above. Hence it is sufficient to consider only one *period*'s equilibrium, albeit describing a range of decisions for the possible *states* that might be observed.

10.2.3 Entry

Entry-deterrence has a literature that is much wider than the applications to electricity markets referred to above (as noted in Chapter 7; the reader is referred to Bagwell and Ramey (1991), and Aghion and Bolton (1987) for surveys). Even in a LRE system, the issue of entry is relevant. The models that follow will show that the spot price varies over time, depending on the cost state of the firm. In high-cost states, the spot price will be high, and potentially attractive to entrants. Furthermore, later chapters will show that there exist incentives for supply firms to "spread" the price, in order to obtain higher contract premiums. With such behaviour, we would expect to see even higher prices in the high-cost states. This is a major reason why regulators encourage potential entrants, in order to discipline the behaviour of generators in high-price periods.

Including a potential entrant requires us to model the decision they face, and the incumbent firms' optimal response to it. The effect of entry on the incumbents' first order conditions is such that a "flat" would be observed in the residual demand curve, at the point at which the entrant would produce, if it chose to enter, and the entrant is assumed to be competitive, offering its capacity at a constant marginal cost. This complication could be solved algorithmically, such as with the grid-search method of Borenstein and Bushnell (1998). However, in a multi-state equilibrium model, an entry decision not only affects the optimality conditions in one state, via its influence on the spot price distribution, it also affects all other states simultaneously. This is further complicated by the entry-decision process, since the potential entrant must assess whether the profits accruing to it in the periods where it would generate are sufficient to cover the fixed costs of entry. Hence we have chosen not to model entry at all, largely because of these analytical difficulties in a multi-state setting. However, we believe it is a fruitful area for future research.

10.2.4 Uncertainty

As discussed above, in LRE, we assume that consumers know the full range of outcomes that occur in the market. More explicitly, we assume that consumers have collected sufficient data to form a discrete distribution of spot price observations, providing a range of prices, each with an associated probability of occurrence. The consumers are unaware of, or unconcerned about, the underlying cause of the stochasticity, or at least they do not consider modelling or predicting it explicitly. From their perspective, they know that, over a given time-frame, any number of price "states" may occur with a given probability, but are unsure as to which of them will, in fact, eventuate in any particular year.

The varying spot prices are, in fact, driven by uncertainty on the supply side of the market. As outlined in Chapter 3, input uncertainty for generation firms is a key component of the risk management triangle. The SDDP model presented in Chapter 8 showed that uncertainty in hydrology is transformed into stochasticity in storage and marginal water values via the reservoir optimisation. While the effect of this on the firm's profit variability was largely negated by the use of storage and market power, it still provided a good degree of spot price variation, and thus potential risk for uncontracted consumers.

Here, we will again reflect the natural variability in hydro systems by including an uncertain stochastic state variable in the firms' cost functions. In this sense, the spot market solutions derived in Section 10.4 may appear to align naturally with the significant, and growing, literature on imperfect competition under uncertainty. However, the treatment of uncertain costs presented here is much simpler than that of Fishelson (1989) or Smeers and de Wolf (1997), for example. In particular, we are assuming risk neutrality on behalf of the supply firms, and, perhaps most importantly, we assume that, in any period, the uncertainty as to which cost state will be observed is resolved before output decisions are made.

The use of cost state is a highly simplified treatment of the hydro reservoir management problem, but one we believe is justified given the aims of this thesis, and the results provided by the empirical analysis in Chapter 8. In many ways, this variable provides a more general treatment of supply input uncertainty, as it can also model non-hydro firms facing uncertainty in fossil fuel prices, for example.

10.3 Statement of the Problem

The general problem facing the generation firm is to find a range of profit maximising generation solutions, one for each cost state, and a single level of contracts that will be sold prior to the resolution of uncertainty in any given period. Given that the uncertainty faced by the generator does not change over time (i.e., the distribution of costs the firm faces is stable), this contract level will be the same throughout the LRE. Hence, in this sense, the optimisation is "time-independent". Thus, in any given period, we can express the problem as:

$$\max_{\bar{g}_i, k_i} \; \mathrm{E}[\Pi_i] = \sum_t \theta^t \left(p\left(G^t\right)\left(g_i^t - k_i\right) - C_i(g_i^t) \right) + f(K)k_i \tag{10.1}$$

where

θ^t is the probability of observing cost state t

p is the spot price, as a function of total generation in state t

g_i^t is firm i's generation level in cost state t, and $\bar{g}_i$ is the vector of generation levels over all states

k_i is firm i's contracting level throughout the LRE

Total Generation $G = \sum_i g_i$; Total contracts sold, $K = \sum_i k_i$

$C_i^t(\;)$ is firm i's total generation cost in state t, as a function of generation and the stochastic cost parameter

$f(K)$ represents the contract price as a function of the number of contracts sold in equilibrium, i.e., the DCC

Problem (10.1) gives rise to a first order condition for each generation level, and the contract level. In this chapter, we do not model the decision maker's consideration of the effect of his/her spot behaviour on the contract negotiation process. As discussed above, spot decisions are made assuming the contract quantity and price are fixed. We also make a symmetric assumption in the contract market - that contract decisions are taken without regard to their effect on future profit maximising generation decisions. This allows us to decompose the spot problem into a deterministic optimisation of the generation variable in each cost state[66], since the generation level in any given state is assumed to not (directly) affect profits in any other state. The contract optimisation problem is simply to find the best level of contracts, given the particular distribution of

[66] The assumption that the cost "state" is revealed prior to the generation level being set seems reasonable, since we are considering the average generation level over time. Hence while the earliest of the short-term generation decisions may be made under uncertainty, as to what sort of year it is going to be, we assume this information will be revealed quickly, and output adjustments made accordingly. Hence the annual average should reflect the true cost state.

spot prices created by the solutions to the spot optimisation. This decision problem is illustrated in Figure 10.3.

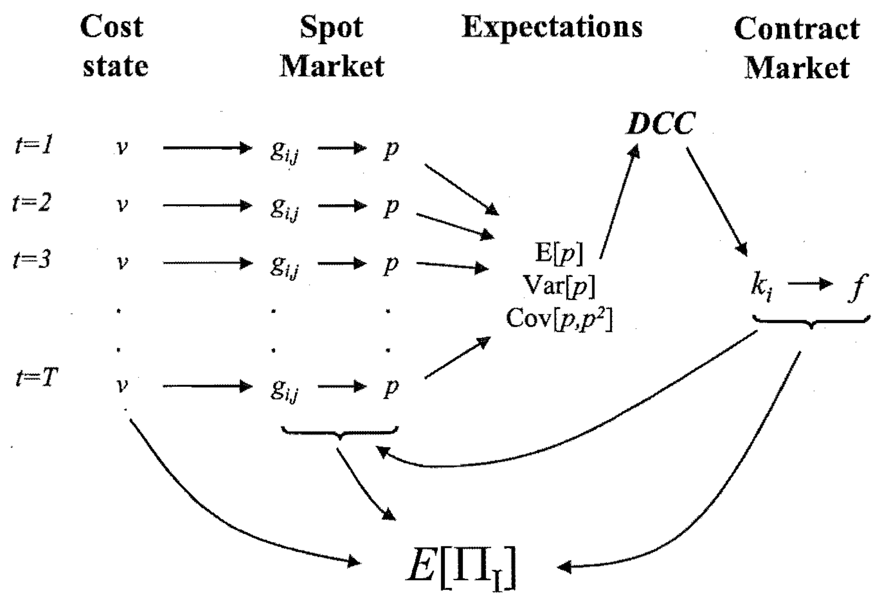


Figure 10.2 Multi-state LRE

Hence, for the "naïve" problem considered in this chapter, the first order conditions will represent

i. The level of contracts that maximises expected profit over all possible cost states, given a fixed level of generation decisions in each cost state.

ii. The short-run profit maximising generation level in each cost state, given a fixed level of contracts, and known cost state, i.e., as though it were a single-state decision under complete certainty.

This framework appears similar to the traditional two-stage framework used by Allaz (1992) and Green (1993), for example. However, since we are searching for long-run joint spot and contract equilibria, these first order conditions must be solved simultaneously. It is important to stress at this point that the formulation of the problem, in this manner, effectively describes an algorithm by which an equilibrium is found.

Assumptions (i) and (ii) above do not reflect ignorance on the part of the decision maker; rather they reflect the fact that the jointly optimal generation and contract decisions will be found by the simultaneous solution of the first order conditions, each of which describe the optimum in the dimension of the solution space corresponding to the particular variable, i.e., with all other variables fixed.

What *does* reflect the naivety of the generation firm, at least at this stage of the model development, is the fact that she ignores the effect that generation decisions will have on the parameters that will in turn *describe* the optimal contract decision. The manipulation of that feedback loop is considered in Chapter 11, but the framework of optimising generation for fixed contracts, and optimising contracts for fixed generation, will be retained.

We will now develop fully the spot and contract market conditions.

10.4 A Model of Spot Market Behaviour

Assume that we have $i = 1..I$ generation firms, and, as argued in Chapter 7, competition in quantities (i.e., Cournot) is the most appropriate representation of the strategic interaction between them. As argued in the previous section, generators maximise short-run profits, ignoring any impact of their output decision on future periods. i.e.,:

$$\max_{g_i^t} \; \Pi_i^t = p\left(G^t\right)\left(g_i^t - k_i\right) + f(K)k_i - C_i^t(g_i^t) \tag{10.2}$$

For the single-state decision, we can assume that generation costs are known with certainty (as the cost state is revealed prior to making the generation decision). As discussed in Chapter 9, generators are assumed to face a linear residual demand for electricity, D, made up of (a) price-responsive and (b) unresponsive load, which is described by:

$$D = \frac{A-p}{b} + L_U \tag{10.3}$$

where A and b are positive parameters describing the load in category (a), and L_U is the total load from category (b), within the given PDC subperiod. Rearranging to form the inverse demand curve (since firms compete in quantities):

$$p(D) = A - b(D - L_U) \tag{10.4}$$

And since we assume that all the electricity offered by generators is taken up by consumers, we can substitute total demand, D with generation, G:

$$p(G^t) = A - b(G^t - L_U) \tag{10.5}$$

Since there is both fixed and variable load in the demand curve, there is an implicit discontinuity in the demand curve at total outputs less than L_U. We assume that total generation is always sufficient to meet the fixed load, and hence the demand curve is always downward sloping and continuous. While Chapter 12 shows that this is always true for the solutions obtained, even if total generation did fall below L_U, we could assume that this represented a response by the retailers to extremely high prices, such as during an electricity crisis. Under Cournot conjectures, where each firm assumes its rivals' output is fixed, we can then state that:

$$\frac{dp}{dg_i^t} = \frac{dp}{dG^t} = -b \tag{10.6}$$

Total costs are quadratic, for each firm, following the general form:

$$C(g_i^t) = g_i^t \left(e_i + \frac{c_i}{2} g_i^t \right)^2 \tag{10.7}$$

so that marginal costs are linear and increasing:

$$\frac{dC(g_i^t)}{dg_i^t} = e_i + c_i g_i^t, \qquad c_i > 0 \tag{10.8}$$

A linear marginal cost function is assumed mainly for analytical convenience - it provides a rough approximation to the stepped nature of the true underlying marginal cost function of a firm with a portfolio of plant with different (constant) marginal costs. But this step function becomes infinitely steep at the point at which generation capacity is exhausted, which is not adequately represented by the linear function. Hence we are implicitly assuming that there are no binding capacity constraints.

As discussed above, a random variable representing the effect of uncertain costs is included. We assume that this is a first-order variability in costs, i.e., it corresponds to a vertical movement in the firm's marginal cost function by scaling the cost intercept e_i. Note that this is equivalent to defining e_i itself as a random variable, but the inter-temporal formulation in Section 10.6 is clearer if a separate variable is used to represent the effect of uncertainty on the firms' costs. In the interests of generality, we will assume that each firm faces a unique cost distribution, and we denote the random scaling variable $\tilde{v}_i$, and assume that $\tilde{v}_i$ has T_i states, denoted v_i^t, $t = 1..T_i$, each with probability of occurrence θ_i^t. Thus (10.8) can be stated generally as:

$$\frac{dC\left(g_i^t\right)}{dg_i^t} = \tilde{v}_i e_i + c_i g_i^t \tag{10.9}$$

Since the generation decision is made once the uncertainty in $\tilde{v}$ has been resolved, it is unnecessary at this stage to make any reference to the distribution of $\tilde{v}$. This is discussed further in Section 10.6, when the complete multi-state model is presented.

In order to find the profit maximising generation level for a given contract level and price, each firm assumes the other's output is fixed and solves the general first order condition to problem (10.2), i.e.,

$$\frac{d\Pi_i^t}{dg_i^t} = g_i^t \frac{dp\left(G^t\right)}{dg_i^t} + p\left(G^t\right) - k_i \frac{dp\left(G^t\right)}{dg_i^t} - \frac{dC_i^t\left(g_i^t\right)}{dg_i^t} = 0 \tag{10.10}$$

Recall that, in this chapter, the spot market decision ignores any effect the generation decision has on the DCC, $f(K)$.

For each state, the firm solves the first order condition (10.10). Since this solution depends on the particular value of $\tilde{v}_i$ that has been observed, we must add a time superscript to the variable representing generation. Substituting in for marginal costs and demand, and simplifying, the first order condition becomes:

$$A-(2b+c_i)g_i^t-b\left(g_{-i}^t-L_U-k_i\right)-v_i^t e_i=0 \tag{10.11}$$

where $g_{-i}^t=\sum_{j\neq i} g_j^t$. The solution to (10.11) is:

$$\hat{g}_i^t=\frac{A-b\left(g_{-i}^t-L_U-k_i\right)-v_i^t e_i}{(2b+c_i)} \tag{10.12}$$

and convexity, expressed as the second order condition, $-(2b+c_i)<0$, is guaranteed for the demand and marginal cost curves assumed above, and, indeed, any reasonable demand and cost curves.

The profit maximising generation level is positive if:

$$A-b\left(g_{-i}^t-L_U-k_i\right)>v_i^t e_i \tag{10.13}$$

The left-hand side of (10.13) is simply the residual demand curve faced by firm *i*, net of the effect of contracts. This condition simply states that as long as some part of the residual demand curve is above the marginal cost curve, a positive generation solution will be found.

As other authors have shown, in the quantity setting framework with reasonable cost and demand curves, a firm's output is decreasing in its rivals':

$$\frac{dg_i^t}{dg_{-i}^t} = -\frac{b}{2b+c_i} < 0 \qquad (10.13a)$$

As its rivals increase their generation and thus depress the market price, the firm decreases output in order to restore the marginal revenue, marginal cost equivalence.

Additionally, the impact that CfD's have on optimal generation is clear:

$$\frac{dg_i^t}{dk_i} = \frac{b}{2b+c_i} > 0 \qquad (10.13b)$$

For every extra unit of electricity sold on forward contract, the firm increases its generation, away from the zero-contracting Cournot output level, by $b/2b+c_i$. As shown in Chapter 7, the decrease in marginal revenue, from increasing generation, only applies to units sold on the spot market, since the contract quantity is sold at a pre-arranged price. This allows a generator with a positive level of contracts to commit to a higher-quantity, lower price output. As c_i tends to zero (i.e., marginal costs tend towards constant), the rate of increase in output tends to its maximum of ½.

Perhaps more interesting is the comparison of (10.12) with the same expression if the firm were to offer the competitive output, in which case (10.10) reduces to the familiar price-marginal cost equivalence relation of perfect competition, i.e., the contract level has no impact on the output level of the firm[67]. Firms would then produce[68]:

$$\hat{g}_{i,PC}^t = \frac{A - b\left(L_U + g_{-i}^t\right) - v_i^t e_i}{c_i + b} \qquad (10.14)$$

[67] Since firms have no influence over the price in the competitive scenario, there is no incentive to move away from P=MC to try and drive up the price on the uncontracted output. This relationship holds true regardless of the level of contracting.

[68] Including the generation of its rival in the PC solution is potentially confusing. This does not imply that the firm responds directly to its competitor's output, but that both firms indirectly determine the market price, and thus the competitive output.

First, as noted by Scott and Read (1996), the firm's contract level no longer has any effect on the profit maximising output. Recall that contracts affect the Cournot solution because a change in the spot price only affects the non-contracted units. However, in a competitive market, firms have no ability to raise or lower the spot price, and hence no incentive to deviate from the zero-contract solution.

Second, it is interesting to note the difference between (10.12) and (10.14) with respect to the stochastic cost variable $v_i^t e_i$. As this varies, the firm exercising its market power alters its output by $1/(2b+c_i)$, for every unit the marginal cost intercept changes. This compares with a generation response of $1/(b+c_i)$ for the firm acting competitively. Clearly, acting competitively leads to a greater variability of generation.

Returning to the Cournot equations, (10.11) defines firm i's optimal generation response to any (fixed) output of the rest of the industry, i.e., equation (10.11), written explicitly for each firm, defines a set of linear reaction functions (Tirole (1988)). The Cournot-Nash equilibrium is found at the intersection of these reaction functions, where no firm has an incentive to alter its output choice, and all firms are acting in optimality with respect to the others' output. Finding the Cournot-Nash equilibrium in each cost state is, of course, equivalent to simultaneously solving the set of first order conditions defined by (10.11), for each state t:

$$\begin{aligned}
&A-(2b+c_1)g_1^t - b\left(L_U + g_{-1}^t - k_1\right) - v_1^t e_1 = 0\\
&A-(2b+c_2)g_2^t - b\left(L_U + g_{-2}^t - k_2\right) - v_2^t e = 0\\
&.\\
&.\\
&A-(2b+c_I)g_I^t - b\left(L_U + g_{-I}^t - k_I\right) - v_I^t e = 0
\end{aligned} \tag{10.15}$$

While these equations can be solved analytically to find profit maximising generation as a function of contracts, we will leave that until after the contract market has been addressed.

10.5 A Model of Contract Market Behaviour

For the purposes of this thesis, the "contract market" will represent the process by which consumers and generation companies negotiate the price and quantity of forward contracts. While the framework presented above indicated that this process occurs prior to the resolution of the cost-state uncertainty for the generators, in order to aid our understanding, we will initially develop the equations under the assumption of perfect foresight on the part of the generators. This assumption will be relaxed when the multi-state model is developed in Section 10.6.

It seems reasonable to assume that if the generation firms are large enough to have a dominant position in the spot market, the same applies in the contract market. We assume that these firms are the only market participants offering contracts for sale. The supply firms face a large number of consumers who wish to purchase contracts. This group of consumers is made up of some load which is unresponsive, and some responsive, to the electricity spot price.

As discussed in Chapter 9, the contracts are standardised and very simple in form. Since the purpose of this thesis is to evaluate the effect of long-term contracting on generation firm behaviour, we ignore load- and time-varying contracts, load interruption provisions, and the many other differentiating factors that might be used by the parties during contract negotiations (for example, consumer loyalty discounts).

We could include multi-year contracts in the model, reflecting the fact that at each contract round, contracts for periods beyond the imminent one could be traded. Normally, this would require some assumption about the effect of time on the contract market (e.g., discounting), and a variable representing the level of contracts, signed prior to the current contract round, for the coming year. However, the characteristics of equilibrium discussed in Section 10.2.2 allow us to generalise the model to account for contracts for any time period. As discussed above, both consumers and generators face an equivalent level of uncertainty for the coming period, at the stage that contracts are signed. If autocorrelation between periods is ignored (which it is), this has two implications. First, generators wish to find a single, optimal contract level (for each LDC subperiod) that will provide them with the greatest level of expected return, given that any of the cost states may occur. Secondly, given that consumers already have a full distribution of the market outcomes that will occur, their DCC will not change from one period to the next. By implication, the consumers face the same degree of uncertainty as to the spot price in *n* time period's time as they do for the coming period. Hence the contract market equilibrium would be identical from period to period, if this aspect of time were included in the model, and the consumers' and generators' respective valuations of the contracts, under uncertainty, will be identical from year to year. Whether an individual period's contract quantity is a combination of previous purchases, or the quantity purchased in this period, is not important to the results that follow.

We assume the supply firms compete in the contract market again according to Cournot conjectures. The quantity-setting model of contract negotiation could be rationalised as follows. Prior to each contract period (which we assume is a regular, standard period of time, e.g., annually), each generator solves the problem described below, and advertises its optimal level of contracts, inviting tenders from the consumers. The consumers bid a uniform contract price for all generators' contracts (according to the aggregate DCC above). At this point, each generator may choose to adjust the quantity made available, having observed their rivals' quantities, and the process starts again until all generators are satisfied with their position in the market. This, of course, corresponds to a Cournot-Nash equilibrium in the contract market.

Other authors have argued that Bertrand competition is a more accurate reflection of the true tendering process that takes place in, for example, England and Wales (e.g., Powell (1993)). Instead of advertising a quantity, supply firms advertise a price. As long as the contract price is greater than each firm's marginal cost, each firm has an incentive to undercut its rival and receive a much larger share of the market. The only stable equilibrium, given the assumptions employed here, is the competitive one, where generators offer contracts at marginal cost, and end up contracted for, and producing, the competitive quantity (Green (1993)). Such fierce competition in the contract market, compared with the relatively uncompetitive spot market seems unlikely[69], as dominant generators will realise that this equilibrium is not, in the long-run, profitable. However, the different degrees of contract, and spot, competition could arise from asymmetries between the markets themselves. It is easy to see how dominant firms might act differently in a market where trading takes place once a year (i.e., the contract market), than in a market that trades every half hour (the spot market).

Green argues that both the time over which contract negotiations are drawn out, and capacity constraints in the spot market will reduce the extent to which a generator who is undercut will lose contract sales, and thus uses Cournot competition as an "outer limit" result. It could also be argued that the degree to which contracts are customised for their purchasers, and the relationship and loyalty that exists between the parties, will also limit lost sales based on contract price alone.

The presence of capacity constraints lends itself to quantity-based contract competition, as firms are unlikely to sell more contracts than their capacity. Thermal plant have well defined capacities, while hydro firms face a degree of uncertainty as to their ability to produce at any stage during the period of the contract, especially during periods of low inflows (e.g., winter). Hence it seems likely that, while we do not model capacity constraints here, quantity-based issues will play a large factor in determining contract equilibria.

Optimal contract levels for generator i are described by the first order condition of (10.2) with respect to the contract level, k_i:

$$\frac{d\Pi_i}{dk_i} = f(K) + k_i\left[\frac{df(K)}{dk_i}\right] - p(G) = 0 \tag{10.16}$$

$f(K)$ is the consumers' aggregate inverse demand for contracts, K, as described by equation 9.32 in Chapter 9:

---

[69] Unless it was regulated

$$f(K) = E[\tilde{p}] + \lambda Var[\tilde{p}]\left(\frac{A}{b} + L_U - K\right) - \frac{\lambda}{4b} Cov\left[\tilde{p}, \tilde{p}^2\right] \tag{10.17}$$

where

$E[\tilde{p}]$, $Var[\tilde{p}]$ and $Cov\left[\tilde{p}, \tilde{p}^2\right]$ are the consumers' estimates of the mean and variance of the spot price, and the covariance of the spot price with its square.

λ is the coefficient of the aggregate consumer risk aversion (Λ in Chapter 9)

A, b and L_U are as defined above

As shown in Chapter 9, the DCC expressed in equation (10.17) may imply a risk premium in addition to the expected spot price for the period covered by the contract, and that this premium is decreasing in the number of contracts purchased by the consumers.

Substituting the spot demand curve, $p(G)$, and contract demand curve $f(K)$ into (10.16), and simplifying, we obtain:

$$E[\tilde{p}] + \lambda Var[\tilde{p}]\left(\frac{A}{b} + L_U - 2k_i - k_{-i}\right) - \frac{\lambda}{4b} Cov\left[\tilde{p}, \tilde{p}^2\right] - \left(A - b(G - L_U)\right) = 0 \tag{10.18}$$

(10.18) can be rearranged to show that the profit maximising level of contracts, assuming fixed generation decisions and rival contracts, is:

$$\hat{k}_i = \frac{1}{2}\left(\frac{A}{b} + L_U - k_{-i}\right) + \frac{1}{2}\left(\frac{\left(E[\tilde{p}] - \frac{\lambda}{4b} Cov\left[\tilde{p}, \tilde{p}^2\right]\right) - \left(A - b(G - L_U)\right)}{\lambda Var[\tilde{p}]}\right) \tag{10.19}$$

Convexity is again guaranteed since the second order derivative, $-4\lambda Var[\tilde{p}]$, is always negative for risk averse ($\lambda > 0$) consumers.

Note that, by implication, we are assuming that consumers use the same value of b in their development of contract demand, as when they purchase electricity in the spot market. This was first proposed in Chapter 9, when we aggregated the responsive consumers' optimal load functions across the industry. Since the contract "round" effectively takes place once, and prior to the spot market period, at that time consumers may have the opportunity to make technological decisions (for example), that will determine their load response in the spot market in each price state. If so, their demand elasticity at the contract market stage may be greater than once contracts are set and the only variation from period to period is in their optimal responses to the spot price. By assuming that the elasticities of responsive consumer load are the same in the spot and contract market, we are suggesting that consumers make no decisions that will influence their ability to respond each period. This could, however, be easily relaxed, by allowing b to take different values in the spot market than in the contract market[70].

The rearranged first order condition above reflects the incentives acting on the firm in the contract market. Two effects, determining the profit maximising quantity of contracts sold by the firm, are evident. First, the firm "anchors" its desired level of commitment to half the maximum possible unhedged consumer load, given its rivals' contract sales. From this, it may increase or decrease contract sales, depending on the sign of the second term in (10.19). The general implication from (10.19) is that the generator has two incentives to sell its output forward:

(i) to receive the risk premium offered by those consumers who are averse to spot variance, and are left unhedged by the firm's rivals' contract sales, and

(ii) to take advantage of inaccurate price expectations on the part of all consumers, or, equivalently, to exploit its own ability to force prices away from the consumers' expectation

If the consumers' expectation of the spot price is significantly less than the Cournot price, the firm may find it optimal to obtain a long position on the forward market, i.e., $\hat{k}_i$ may be negative. The intuition of this situation is relatively clear - if contracts are cheap enough, the firm would do well to purchase contracts (off its rivals) at a cheap price, and

[70] This could be complicated if the elasticity achieved in the spot trading periods was a function of the elasticity at the contract market stage.

sell the generation back to the spot market at the favourable Cournot price, thus making a profit on the contracted units.

It is expedient to note, at this stage, that any interpretation of the first-order conditions, either for spot or contract decisions, must be made in light of what these equations actually describe, namely, a single-dimension optimum in a multi-variable problem. To illustrate, one could find it intuitively surprising that the first order condition for contracts does not contain any terms relating to the cost function, which would seem to have a significant effect on the optimal contract decision. However, recall that the joint consideration, and indeed optimisation, of contracts and generation (and thus costs) is made by solving all first-order conditions simultaneously. Hence the first order conditions developed above describe a "partial equilibrium", rather than unilaterally representing all the factors determining the globally optimal choice of contracts and generation.

Again, each firm solves the first order condition. Thus, simultaneously solving the set of equations described by (10.18), results in a Cournot-Nash contract market equilibrium.

We will now use the optimality equations developed for the spot and contract market in a multi-state context.

10.6 A Multi-state Model of Joint Equilibria

10.6.1 Joint Spot-Contract Equilibrium

As already discussed, each of the Cournot spot and contract equilibria can be found by solving the set of optimality equations developed above in the relevant section. In order to find a joint equilibrium between the markets, however, we need to define a system containing both spot and contract equations.

At the simplest level, a single-state, joint spot-contract equilibrium for two Cournot firms (i.e., a duopoly) can be found by solving the system of the two spot market, and two contract market equations, simultaneously, for the generation and contract quantity for each firm (Equation (10.20)). If the expected spot price, variance of the spot price and stochastic cost parameter are all assumed to be fixed and exogenously determined, all four equations in this system are linear, and thus a solution to the four decision variables could be found (providing the determinant is non-zero) by standard linear algebra.

$$
\begin{aligned}
&A-(2b+c_1)g_1-b(g_2+L_U)+bk_1-v_1e_1 &&=0\\
&A-(2b+c_2)g_2-b(g_1+L_U)+bk_2-v_2e_2 &&=0\\
&E[\tilde{p}]+\lambda Var[\tilde{p}]\left(\frac{A}{b}+L_U-2k_1-k_2\right)-\frac{\lambda}{4b}Cov(\tilde{p},\tilde{p}^2)-A+b(G-L_U) &&=0\\
&E[\tilde{p}]+\lambda Var[\tilde{p}]\left(\frac{A}{b}+L_U-2k_2-k_1\right)-\frac{\lambda}{4b}Cov(\tilde{p},\tilde{p}^2)-A+b(G-L_U) &&=0
\end{aligned}
\tag{10.20}
$$

As the number of firms, and number of cost states, increase, so do the number of equations in this system. However, another relationship between the equations must be accounted for as we extend beyond one state. In the linear system of equations (10.20), we have ignored the effect of variability in the cost state $\tilde{v}$ on the contract market. As more states are introduced, a variety of spot market equilibria are found, thus determining the consumers' estimates of mean, variance, and covariance, used to form the DCC.

In order to achieve the goals of the LRE framework, described in Section 10.2, two important enhancements must be made. These are:

- Acknowledge that the generators do not know, at the time of the contract negotiation process, what spot market equilibrium will be realised in the period covered by the contract.

- Explicitly develop equations for the estimates of the mean, variance and covariance of the spot price, based on the range of spot price outcomes driven by the variability in $\tilde{v}$

Both of these enhancements require us to reconstruct system (10.20) so that, in equilibrium, all possible spot market conditions are solved simultaneously with the contract market optimality conditions (10.18). The total number of spot market equilibria which must be solved by this system is entirely dependent on the assumptions made about the distributions of the cost state.

10.6.2 Spot Market Equilibria

We assume that the probability distribution of $\tilde{v}$, whether it is known analytically or numerically, is discretised into a finite number of states. Each of the supply firms will face T_i different cost states, each with fixed probability θ_i^t, and a unique profit-maximising equation (10.11) for each state.

Each spot market equilibrium is the solution to a particular combination of these equations (one equation for each firm) and thus the total number of spot market equilibria is determined by the total number of possible combinations of states. This, in turn, is determined by the assumptions made about the relationship between the two firms' cost state distributions, in particular, the implications for the joint distribution of the two firms' cost states.

An example may be helpful. Consider the simple case where both firms face an identical distribution of cost states, which are equally probable, and we will refer to these states as low, medium and high costs (L, M, H). Each state has a $^1/_3$ probability of occurring.

It is conceivable that the correlation between these firms' states is very high if, for example, both firms generate using hydro, and have reservoirs in the same region, and are thus subject to the same weather patterns. If $\tilde{v}$ is to be interpreted as an approximation of the effect of variable inflows on marginal water values, then it seems broadly reasonable that both firms will simultaneously face low, medium or high costs. In fact, if both firms were thermal, the degree of correlation between cost movements could be even greater, since both firms would be tied to the same world fuel price. In such a case, we would observe the following joint distribution:

		Firm 1		
		L	M	H
Firm 2	L	1/3		
	M		1/3	
	H			1/3

Thus we would only have 3 possible spot market equilibria, corresponding to both firms being in either the low, medium or high cost state. This is the minimum number of total possible states, given these particular individual cost distributions.

The same number of total possible spot market equilibria occur if the states are, for some reason, perfectly negatively correlated, i.e.,

		Firm 1		
		L	M	H
Firm 2	L			$^1/_3$
	M		$^1/_3$	
	H	$^1/_3$		

At the other extreme is the scenario where, for a given firm in a particular state, its rival may be in any of its three states. This would be an appropriate way of modelling a market that consisted of one hydro and one thermal firm, for example. The hydro firm's marginal water values are driven by meteorological factors, as above, and the thermal firm's marginal costs are significantly influenced by the price of fossil fuels. It is reasonable to assume that the two driving factors of cost variations are independent. In such a case, the joint distribution could be:

		Firm 1		
		L	M	H
Firm 2	L	$^1/_9$	$^1/_9$	$^1/_9$
	M	$^1/_9$	$^1/_9$	$^1/_9$
	H	$^1/_9$	$^1/_9$	$^1/_9$

Thus we have nine possible spot market equilibria, corresponding to 9 combinations of the firms' first order conditions, each with associated joint probability $^1/_9$.

We have, in fact, defined the conditional probabilities of a firm's rival being in state t_j, given that the firm itself is in t_i. Each element of the above tables is simply the product of two individual probabilities, according to the following familiar statistical result:

$$P(t_i \cap t_j) = P(t_i \mid t_j).P(t_j) = P(t_j \mid t_i).P(t_i)$$

In the highly correlated case, we implicitly assumed that the conditional probability, $P(t_i \mid t_j) = 1$. In the second case, when the states are independent, the individual probabilities are simply multiplied, i.e., $P(t_i \mid t_j) = P(t_j)$.

Clearly, this can be generalised further to allow any relationship between the firms' cost state distributions. The two cases discussed above are convenient, since the total number of state combinations, and thus spot market equilibria, is an intuitively easy calculation. In the perfectly correlated case, where all conditional probabilities are 1, and both firms face the same number of states we would have $\overline{T} = T_1 = T_2$ pairs of equations, recalling that T_i is the total number of states faced by firm i. In the totally independent state case, we would have $\overline{T} = T_1 T_2$ total equilibria.

In between these two extremes, the number of possible spot equilibria, $\overline{T}$, is less obvious. Consider the case where, for example, the conditional probability table was:

		P(Firm 1|Firm 2)		
		L	M	H
Firm 2	L	$^1/_3$	$^1/_3$	$^1/_3$
	M	$^1/_2$	$^1/_2$	0
	H	0	0	1

Then, multiplying each conditional probability by the chance that Firm 1 will be in that state, we obtain the joint distribution table, $P(t_1 \cap t_2)$:

		Firm 1		
		L	M	H
Firm 2	L	$^1/_9$	$^1/_9$	$^1/_9$
	M	$^1/_6$	$^1/_6$	0
	H	0	0	$^1/_3$

Thus the total number of states is 6. Hence we can state the general result that the total number of equilibria that can be observed in the market is the number of non-zero joint probabilities (or conditional probabilities). Note that even if this number is equal to T_1T_2, this does not imply independence. Only if all joint probabilities are nonzero and the rows of the conditional or joint probability matrix are identical (implying that the conditional probabilities are not dependent on which state the firm is in) do we have state independence between firms.

It is easy to see how this model could be further extended to account for scenarios where the information is asymmetric between firms. For example, one firm could have better information about which state its rival will be in, than the other firm. Alternatively, a thermal firm may be certain of its fuel cost prior to contract negotiations, while its rival is still unaware, and uncertain of its own costs. These situations would complicate the models that follow, as the equilibria expected by the firms, when assessing their contract strategy, would not necessarily be the same as those experienced once uncertainty is revealed. Hence we leave these extensions as an area for future work.

Given the distribution of cost states for each firm, and the conditional probabilities, we know that for every state that has a positive conditional probability, a spot market equilibrium, corresponding to the firms' respective states, is possible. We can now form the full system of spot market optimality conditions, and restate the contract problem for each firm by developing explicit expressions for the mean, variance and covariance terms of the DCC.

10.6.3 Contract Optimisation Under Uncertainty

In order to reflect the generator's uncertainty of its own, and its rival's, costs, at the time of contract negotiations, each generator's contract market problem is stated as:

$$\max_{k_i} \; E\left[\tilde{\Pi}_i\right] = f(K)k_i + E\left[p\left(\tilde{G}\right)\left(g_i - k_i\right) - \tilde{C}_i(g_i)\right] \tag{10.21}$$

Equation (10.21) yields a first order condition with respect to k_i of:

$$\frac{d\Pi_i}{dk_i} = f(K) + k_i\left[\frac{df(K)}{dk_i}\right] - E\left[p(\tilde{G})\right] = 0 \tag{10.22}$$

Substituting for the DCC, we obtain:

$$E[\tilde{p}] + \lambda Var[\tilde{p}]\left(\frac{A}{b} + L_U - 2k_i - k_j\right) - \frac{\lambda}{4b}Cov\left[\tilde{p}, \tilde{p}^2\right] - E\left[A - b\left(\tilde{G} - L_U\right)\right] = 0 \tag{10.23}$$

Under the assumptions of the LRE framework, the consumers have observed, or are aware of, all possible spot market equilibria that may occur. At the time contracts are negotiated, generators have no better predictions of the spot price than the consumers do, since the cost state for the contract period has not yet been revealed. Thus the generators cannot profit from a difference between expectations (as was suggested initially in Section 10.5). Hence the consumers' and generators' expectations of the spot price are equal, and the first and last terms of (10.23) cancel[71]. This provides the firm with a simpler first order condition of:

$$\lambda Var\left[\tilde{p}\right]\left(\frac{A}{b} + L_U - 2k_i - k_j\right) - \frac{\lambda}{4b}Cov\left[\tilde{p}, \tilde{p}^2\right] = 0 \tag{10.24}$$

The optimal level of contracts is now:

$$\hat{k}_i = \frac{1}{2}\left(\left(\frac{A}{b} + L_U - k_j\right) - \frac{Cov\left[\tilde{p}, \tilde{p}^2\right]}{4bVar\left[\tilde{p}\right]}\right) \tag{10.25}$$

(10.25) shows that each generator will now always sell contracts equal to less than half the maximum unhedged load, since we know that with positive prices, the covariance term is always positive.

---

[71] A hedging equation based entirely on the variance of wealth, rather than the expected level of wealth, would also appear to ignore the fact that perceptions of risk may be determined by the ratio of wealth variance to mean wealth, i..e, the coefficient of variation is assumed to be constant for any level of variance. This is a result of the mean-variance utility, which separates these two aspects of the wealth distribution.

While the average spot price may not be of concern to the firm, we must develop full expressions for the remaining variance and covariance terms, which provide the link between the spot market behaviour and the contract market equilibrium. As discussed above, the expressions for these terms depend on the assumptions made about the correlation between cost states. We will develop the necessary equations for the independent and perfectly correlated case, separately.

Expressions for $E[\tilde{p}]$, $Var[\tilde{p}]$ and $Cov\left[\tilde{p},\tilde{p}^2\right]$

As outlined above, the number of spot market equilibria that can be observed by the firm, when developing their estimates of mean and variance, is determined by the number of non-zero conditional probabilities. We also showed that, using a familiar statistical result, we can develop the probabilities of these equilibria being observed, i.e., the joint probabilities.

Let $\theta_i^{s|r}$ be the conditional probability that firm *j* is in state $s \in \{T_i\}$, given that firm *i* is in state $r \in \{T_j\}$. The spot price resulting from states *r* and *s* being observed for the respective firms, $\hat{p}^{r,s} = A - b\left(\hat{g}_i^r + \hat{g}_j^s\right)$ would occur with probability $\theta_i^r \theta_i^{s|r}$. $\hat{g}_i^t$ is the Cournot-Nash equilibrium generation solution for firm *i* in state *t*.

In such a case, with $\hat{p}^{r,s}$ defined as above, the expressions for the mean, variance and covariance terms become:

$$E[\tilde{p}] = \sum_{r \in T_i} \sum_{s \in T_j} \theta_i^r \theta_i^{s|r} \left(\hat{p}^{r,s}\right) \tag{10.26}$$

$$Var[\tilde{p}] = \sum_{r \in T_i} \sum_{s \in T_j} \theta_i^r \theta_i^{s|r} \left(\hat{p}^{r,s} - E[\tilde{p}]\right)^2 \tag{10.27}$$

Since, for every combination of *r* and *s*, $\hat{p}^{r,s}$ and $\left(\hat{p}^{r,s}\right)^2$ co-occur with probability 1, the covariance can be expressed as

$$
\begin{aligned}
Cov\left[\tilde{p},\tilde{p}^2\right] &= \sum_{r\in T_i}\sum_{s\in T_j}\theta_i^r\theta_i^{s|r}\hat{p}^{r,s}\left(\hat{p}^{r,s}\right)^2-\left(E[\tilde{p}]\right)\left(E\left[\tilde{p}^2\right]\right) \\
&= \sum_{r\in T_i}\sum_{s\in T_j}\theta_i^r\theta_i^{s|r}\left(\hat{p}^{r,s}\right)^3-\left(E[\tilde{p}]\right)\left(\sum_{r\in T_i}\sum_{s\in T_j}\theta_i^r\theta_j^{s|r}\left(\hat{p}^{r,s}\right)^2\right)
\end{aligned}
\tag{10.28}
$$

For states that are independent, $\theta_i^{s|r}=\theta_j^s$, while for states that are perfectly correlated, $\theta_i^{s|r}=1$, which provides a much simpler set of expressions (and a much lower number of total equations).

Inserting equations (10.26)-(10.28) into the contract market optimality conditions results in a complete, two-firm LRE system of $2\overline{T}+2$ equations.

10.6.4 Solution to the Full LRE System

We intend to find solutions to the LRE system under a range of assumptions regarding the interactions between firms' cost states. Due to the explicit expressions for variance and covariance, the contract equations are non-linear, making analytical solutions to the full system intractable with any reasonable number of cost states (> 2). Thus we will pursue numerical solutions only, which are reported in Chapter 12.

10.7 Conclusions

This chapter has proposed, and modelled, a framework for analysing the dynamics that exist within, and, more importantly, between the contract and spot market for electricity. Following the theme of earlier chapters, a model was presented that incorporated dominant supply firms that faced uncertain costs, and a contract-buying demand side that was a mix of load-responsive and load-unresponsive consumers.

A key element of this framework was that the market was in a state of long-run equilibrium, immune to the uncertainties of entry, regulation and demand growth. It was assumed that consumers had already seen a large enough sample of supply firm behaviour to be sure that any future price would belong to the distribution formed from this sample. Thus, generators had no incentive to "surprise" consumers, by acting in a manner not seen before in the market. Thus the solution to the model would represent optimal contract and generation policy that would be stable throughout time, as long as the market was in this state of long-run equilibrium.

The model essentially described how this distribution would come about, by describing the profit-maximising behaviour, *á la* Cournot, of the supply firms in the spot market, for a fixed level of contracts. Equivalently, the profit maximising contract decision was described, for a given level of generation in each state. Since the spot equilibrium depended on the (uncertain) cost state each firm experienced, this produced a range of equilibria, and thus spot prices, which the consumers observed. Once the consumers had formed expectations of the mean, variance and covariance of the spot price, a contract market equilibrium could be found, again, following Cournot conjectures for the firms selling contracts. The linked nature of the spot and contract markets was modelled by simultaneously solving the system of optimality conditions for each market, a process by which the effect of the optimal contract decision on the optimal generation decision, and vice versa, is modelled implicitly. This was a complex non-linear system, and hence did not lend itself to traditional linear algebra solution methods.

However, this model was "naïve", in the sense that while the firms were aware that their varying responses to cost uncertainty drove the spot market distribution used by consumers, thus determining contract prices, they did not act to deliberately influence this process. Instead, the firms acted in short-run optimality, and let the spot price distribution be determined "by default". While the profit accruing to such behaviour may be attractive (compared to perfect competition), the question naturally remains as to whether a supply firm could increase profits even further, by altering their behaviour in the spot market to deliberately influence the formation of contract prices.

The next chapter addresses the question of whether an equilibrium can be found between the two firms in the spot market, under Cournot conjectures, that supports deliberate influence of the DCC.

11

MARKET DESTABILISATION

11.1 Introduction

In the last chapter, a model was developed that described conditions for spot and contract market equilibria, for dominant electricity suppliers. We outlined the dynamics of a market that is in a state of "long-run equilibrium" (LRE) where consumers have full knowledge of the distribution of spot prices that can occur in the market. This distribution was created by the dominant supply firms responding to variability in marginal costs, in a Cournot, short-run profit maximising manner.

Previously, we have shown that, for a dominant firm, the profit-maximising response to a *fixed* level of contracts is for the firm to increase output above, and thus depress the spot price below, the profit maximising solution in the absence of contracts. We showed that, in the absence of risk premiums (i.e., contracts are valued at the expected spot price), contracts decreased total profit for these generators. Few other incentives existed for risk-neutral firms to sell contracts (for example, avoiding regulation, or being forced to by regulation). However, the results of Chapter 10 showed that from a somewhat naïve perspective, when certain forms of risk premiums are offered by consumers, it is optimal for generators to sell a non-zero quantity of contracts.

Given that these risk premiums were a result of the variance created by the generators responding to cost variability in a short-run profit maximising way, the question naturally arises as to whether profits could be further increased by the firm taking a more intentional approach to affecting long run profit, i.e., contract prices. Rather than let the distribution of spot prices be determined by a short-run profit maximising response to cost variability, spot market equilibria could be manipulated to increase contract market profits. If the increase in contract profits outweighs the sacrificed short-run spot profits, this would be an attractive joint spot-contract strategy for the firms.

While the effect of a firm's hedging position on spot market behaviour is relatively well understood in the literature, the concept of dominant firms adopting a spot market strategy designed to advantage their contract negotiations is a largely unexplored research area. Section 11.2 outlines the work that has taken place in this area already, and develops the intention of this chapter more fully. Section 11.3 re-introduces the LRE model, and enhances it with the new incentives. Concluding remarks are made in Section 11.6.

11.2 Background

In order to investigate how a firm's spot market strategy can be used to increase contract profit, the ways in which a firm's spot market behaviour influence contract prices must be examined. If we ignore the possibility that the firms change their form of strategic interaction, and assume that they continue to follow Cournot conjectures[72], then the only avenue for firms to influence the contract market is through the consumers' demand curve for contracts. Since the contract price the generators receive is in part already determined by the Cournot equilibrium quantity of contracts they sell, any improvement in contract prices can only come from shifting the position of the DCC.

Recall the form of the DCC that represents our model of the contract consumers:

[72] Powell (1993) and Allaz (1992) investigated the case where firms colluded in the contract market.

$$f(K) = E[\tilde{p}] + \lambda Var[\tilde{p}]\left(\frac{A}{b} + L_U - K\right) - \frac{\lambda}{4b} Cov\left[\tilde{p}, \tilde{p}^2\right] \tag{11.1}$$

We have assumed that generators have no ability to influence consumers' demand curves for electricity, or their risk aversion. Hence the only mechanism that generators can use to influence the position of the DCC is the distribution of the spot prices, hence the mean, variance and covariance terms that appear in equation (11.1).

It is clear from this equation that the contract price can be improved by increasing the expected mean spot price, and/or the variance of the spot price, and/or decreasing the covariance of the spot price with its square. While these strategies will be addressed analytically below, we will first outline the relatively few studies that have investigated such strategies.

Green (1993) proposed an infinitely repeated gaming model where consumers used the previous period's mean spot price as a perfect predictor of the mean spot price in the next period, i.e., a lag-1 prediction. A single spot and contract decision was made each period, and no risk aversion modelled, so that the effect of increasing price expectations could be isolated. Generators chose their output quantities in a given period not only considering the effect it would have on this period's spot price and spot profitability, but also the indirect (discounted) effect it would have on the next period's contracting demand and thus contract profitability. The influence of potential future contract profits acted in opposition to the traditional result that contracts cause a dominant firm to increase output, making contracts an attractive regulation tool. In fact, Green's conclusion showed that the effect of this strategy was quite dramatic:

> *"...the presence of the contract market has made very little difference to the 'worst-case' [Cournot] outcome. If the generators believe that they can affect future contract prices by bidding up the present pool price, they could raise the pool price to near the [zero-contracting] Cournot level, while remaining fully contracted"* Green, (1993) p11

Since consumers were risk neutral in Green's multi-period model, there was no need to address the issue of spot price variance. In fact, very little treatment of supply firms influencing spot price volatility is evident in the literature. However, the concept of spot volatility as a result of dominant firm behaviour is not foreign to the large body of work in the electricity area. Mount (2001) and Guan, Ho and Pepyne (2001) explicitly addressed the issue of price "spikes" in electricity markets, and many other authors have agreed that the structure of electricity markets is such that even small firms can possess the power to cause large price spikes, by withholding capacity from the market at times of high, and inelastic, demand. Mount argued that this volatility, even in the absence of market power, was a result of the form of the market-clearing auction (i.e., all successful sellers are paid the same price). Under the uniform price auction, the aggregate supply curve tends to be significantly steeper, and thus relatively price inelastic, causing even small errors in forecasted load to be amplified into high price volatility. Mount showed that if a paid-as-offered auction, or discriminatory price auction, were adopted, the aggregate supply curve would be more price elastic, reducing the effect of forecast errors on price volatility.

However, all of these analyses have assumed that such behaviour is driven by the opportunity for generators to make short-term profits. This is appropriate for markets that are tightly constrained, and the spikes are (usually) the result of firms' capitalising on high and inelastic loads which are usually of a short duration, often lasting less than an hour.

Only Bunn, Larsen and Dyner (1997) have suggested that generators may attempt to amplify price volatility in order to increase long-run profits through contract prices. The authors provide a system dynamics model of an electricity market, reviewed in Chapter 7, which investigates three volatility-increasing strategies, each one a variant on a regular capacity withdrawal theme. Their results clearly indicated that contract profits could be increased if generators adopted these strategies. Whether or not the authors' "regular spiking" strategies are sustainable in environments where dominant firm behaviour is watched closely by a market regulator, remains to be seen. However, Bunn did note that even a regulator would find it difficult to detect this behaviour during naturally high demand/low supply periods:

> *"As volatility is already high during such times, the potential for even a small amount of tactical behaviour of the sort indicated in this*

> *article will be very hard to notice, but potentially quite profitable."*
> Bunn, Larsen and Dyner (1997), p286

While the analysis of Bunn *et al* reflected the state of the England & Wales market, this is also true of a hydro dominated system such as New Zealand, where spot price volatility is expected to occur as a natural result of hydrological variation. However, it is not the intention of this thesis to show that such destabilisation strategies are being employed in any market, or to prove otherwise. Rather, we wish to investigate whether firms have profit-making incentives to deliberately influence spot price volatility, and the spot price mean, and if such strategies form a stable Cournot equilibrium in both the spot and contract market.

The structure of the problem, as defined here, is for each firm to select generation quantities in each of a finite number of cost states, such that total expected profit is maximised, once the optimal contract strategy is chosen. This seems very similar, in form, to a nonzero-sum stochastic game with a finite state space. These games, initially analysed by Shapley (1953), proceed through a number of stages, and at each stage, players select a decision from a finite set of alternatives. The transition to the next stage is determined by a probability distribution. Thus stochastic games are very similar to Markov Decision Processes (MDPs). In many types of these games, stationary strategies, such as the policies found for MDPs, are shown to be optimal. Other settings give rise to randomised strategies, where a player randomly selects a strategy, from a set of alternatives, at each stage. In stochastic games, concepts such as Nash equilibria are much more difficult to define. We did not pursue this option as a potential way of representing the situation modelled here, but some of the results in Chapter 12 motivate it as an alternative way of finding "destabilisation" strategies in more general settings.

11.3 A Model of Market Destabilisation

We now proceed to an analytical model of contract-price influence. To support the concept of LRE, we again assume that consumers have full access to the range of spot prices that can occur in the market, and form their DCC in exactly the same fashion as that used for the "naïve" model developed in Chapter 10. Also, the random cost state, $\tilde{v}$, has the same properties, and effect on costs, as in that chapter. In order to simplify the equations presented here, we shall assume that the firms face the same number of perfectly correlated cost states, so that for each state that firm i finds itself in, there is only one possible state for firm j. The extension to other, more complicated cost state relationships is relatively straight-forward, as shown in Chapter 10, but would unnecessarily complicate the equations developed below. The effect of cost-state relationships on the results presented below is discussed in Section 11.3.3, and both cases will be solved for, numerically, in Chapter 12.

The generators, again, face the following problem:

$$\begin{aligned}\max_{g,k} E\left[\Pi_i\right] &= \sum_{t=1..T_i} \theta_i^t \Pi_i^t \\ &= \sum_{t=1..T} \theta_i^t \left(p\left(G^t\right)\left(g_i^t - k_i\right) + k_i f - C_i^t\left(g_i^t\right)\right)\end{aligned} \tag{11.2}$$

As will be shown below, the first order generation conditions for problem (11.2) may not have a solution in some states, for certain combinations of the parameters involved. This is largely a result of the fact that the profit function (11.2) is not strictly concave in output, when generators consider the effect their spot behaviour has on contract prices. This is addressed in Section 11.3.4, and leads us to redefine the variables measuring risk, so that spot market equilibria can be found.

11.3.1 Contract Market Decisions

The contract problem is unchanged from the naïve model. The contract decision still amounts to choosing the best quantity of contracts to sell, given the consumers' estimates of variance and covariance, and thus their DCC. Hence the first- and second-order conditions with respect to contracts are identical to those developed in the last chapter, i.e.,

$$\lambda Var\left[\tilde{p}\right]\left(\frac{A}{b} + L_U - 2k_i - k_j\right) - \frac{\lambda}{4b} Cov\left[\tilde{p}, \tilde{p}^2\right] = 0 \tag{11.3}$$

$$-\lambda Var\left[\tilde{p}\right] < 0 \tag{11.4}$$

Again, (11.4) always holds, guaranteeing that any solution to (11.3) maximises profit, and we can substitute the full expressions for variance and covariance into equation (11.3).

11.3.2 Spot Market Optimality Conditions

In the naïve case, the generation decision made in an individual state (corresponding to a realisation of the cost variable) was assumed to only have impact in the particular state. Given that generators now wish to take advantage of the influence of equilibrium spot prices on contract prices, the optimality condition on generation, in a particular state, must reflect the maximisation of both spot *and* contract profits, the latter affecting profit in all states:

$$\frac{dE[\Pi_i]}{dg_i^r} = 0 = \theta_i^r \frac{d\Pi_i^r}{dg_i^r} + \sum_{t \neq r} \theta_i^t \frac{d\Pi_i^t}{dg_i^r} \tag{11.5}$$

Note that the effect of generation on the current state's profits is now weighted by the probability of that state actually occurring. This is to be expected, as we are now trading off profits within the state, with those made over all states. It is these profits that are represented by the second derivative term in (11.5). Initially, it may seem natural to include discounting at this point. However, the simplicity of equation (11.5) disguises an important aspect of our model that was initially raised in Chapter 10. A model of discounting would require that, in each period, the problem facing the generator is to maximise the current period's profits, plus discounted future profits. This would imply that the decision process proceeds through time. However, in the equilibrium model used here, it is as though the decision maker is "outside" of time, since the distribution of states is constant through time. While this approach allows us to make use of simplifying assumptions (such as the stability of expectations), it makes the inclusion of discounting irrelevant. Aspects such as discounting would become relevant if we wished to model a changing distribution of costs through time (such as a regime-switching model representing years of drought), and would be better modelled using a Stochastic Dynamic Program or Markov Decision Problem. However, these approaches were rejected for this phase of the analysis, since they were too analytically complicated to model the complex interactions between the spot market, contract market, and the impact on the distribution of prices observed by consumers.

Since we know that the generation level chosen in the current state only affects contract revenue in other states, and not (directly) the generation in any other state, we can use equation (11.2) to restate (11.5):

$$
\begin{aligned}
0 &= \theta_i^r\left[g_i^r\frac{dp(G^r)}{dg_i^r}+p(G^r)+k_i\left(\frac{df(K)}{dg_i^r}-\frac{dp(G^r)}{dg_i^r}\right)-\frac{dC_i^r(g_i^r)}{dg_i^r}\right]+(1-\theta_i^r)\left(k_i\frac{df(K)}{dg_i^r}\right)\\
&= \theta_i^r\left[g_i^r\frac{dp(G^r)}{dg_i^r}+p(G^r)-k_i\left(\frac{dp(G^r)}{dg_i^r}\right)-\frac{dC_i^r(g_i^r)}{dg_i^r}\right]+k_i\frac{df(K)}{dg_i^r}
\end{aligned}
\tag{11.6}
$$

Again, we have a system of $2\overline{T}$ spot equations, where $\overline{T}$, in the correlated state case, is the total number of states that can possibly be observed. Each spot equation is paired with its rival's, for each state, in order to find an equilibrium (see Section 11.3.4). These equations are again combined with the pair of contract market optimality conditions, to form a complete system of $2\overline{T}+2$ equations.

Substituting in the equations describing demand and marginal cost (10.4 and 10.8), the expression inside the square brackets in (11.6) is identical to the first order condition in the naïve model. We now must determine the derivative of the DCC with respect to an individual generation decision.

Using the chain rule, the derivative of the DCC, with respect to generation, can be written as:

$$
\frac{df(K)}{dg_i^r}=\frac{df(K)}{dE[p]}\frac{dE[p]}{dp^r}\frac{dp^r}{dg_i^r}+\frac{df(K)}{dVar[p]}\frac{dVar[p]}{dp^r}\frac{dp^r}{dg_i^r}+\frac{df(K)}{dCov(.)}\frac{dCov(.)}{dp^r}\frac{dp^r}{dg_i^r}
\tag{11.7}
$$

The majority of the terms in (11.7) are relatively easy to express, however, the derivatives of variance and covariance require a more detailed development. Recall that we are assuming perfect correlation between firms' cost states, and so:

$$E[\tilde{p}] = \sum_t \theta^t \left(p^t\right) \tag{11.8}$$

$$Var[\tilde{p}] = \sum_t \theta_t \left(p_t - E[\tilde{p}]\right)^2 \tag{11.9}$$

$$Cov\left[\tilde{p}, \tilde{p}^2\right] = \sum_t \theta^t \left(p^t - E[\tilde{p}]\right)\left(\left(p^t\right)^2 - E\left[\tilde{p}^2\right]\right) \tag{11.10}$$

The derivative of the average price with respect to an individual price observation is simply its probability of occurrence, θ_i^r. Then, using the composite function derivative law:

$$\begin{aligned}\frac{dVar[p]}{dp^r} &= 2\theta_i^r \left(p^r - E[p]\right)\left(1 - \theta_i^r\right) - \theta_i^r \sum_{t \neq r} 2\theta_i^t \left(p^t - E[p]\right)\\ &= 2\theta_i^r \left(p^r - E[p]\right) - \theta_i^r \sum_t 2\theta_i^t \left(p^t - E[p]\right)\end{aligned}$$

Since all probabilities, θ_i^t sum to unity:

$$\begin{aligned}\frac{dVar[p]}{dp^r} &= 2\theta_i^r \left(p^r - E[p]\right) - \theta_i^r \left(2E[p] - 2E[p]\right)\\ &= 2\theta_i^r \left(p^r - E[p]\right)\end{aligned} \tag{11.11}$$

And using the product derivative law for the derivative of covariance:

$$\frac{dCov(p, p^2)}{dp^r} = 3\theta_i^r \left(p^r\right)^2 - \theta_i^r E\left[p^2\right] - 2\theta_i^r p^r E[p] \tag{11.12}$$

We can now form the full expression for equation (11.7). Recall that the contract demand curve, i.e., the contract price as a function of contract level, variance and covariance, is given by 11.1:

$$f(K) = E[\tilde{p}] + \lambda Var[\tilde{p}]\left(\frac{A}{b} + L_U - K\right) - \frac{\lambda}{4b} Cov\left[\tilde{p}, \tilde{p}^2\right] \tag{11.13}$$

Thus, using equations 11.7 – 11.12, the derivative of the contract demand curve with respect to generation is:

$$\frac{df(K)}{dg_i^r} = \theta_i^r \left[-b - \lambda b \left(\frac{A}{b} + L_U - \sum_i k_i \right) \left(2\left(p^r - E[p]\right) \right) + \frac{\lambda}{4} \left(3\left(p^r\right)^2 - E\left[p^2\right] - 2p^r E[p] \right) \right] \tag{11.14}$$

where $p^r = A - b\left(g_i^r + g_j^r - L_U\right)$. Now, substituting (11.14) into the generation first order condition (11.6) yields:

$$\begin{aligned} 0 = \theta_i^r \left[A - b\left(2g_i^r + g_j^r - L_U\right) + bk_i - \left(v_i^r e_i + c_i g_i^r\right) \right] + \theta_i^r k_i \Bigg[-b - \\ b\lambda \left[2\left(\frac{A}{b} + L_U - \sum_i k_i \right) \left(p^r - E[p]\right) - \frac{1}{4b} \left(3\left(p^r\right)^2 - E\left[p^2\right] - 2p^r E[p] \right) \right] \Bigg] \end{aligned} \tag{11.6a}$$

Note that, again, we assume that consumers exhibit the same demand elasticity at the contract stage as they do at the spot "stage", thus all terms relating to the spot demand curve are equal.

We can see that the effect of a state's output on the mean, variance and covariance terms is weighted by the frequency with which the state occurs, and the effect on the states short-run spot profits is also weighted by θ_i^r, representing the effect on total expected profit. So while behaviour in low-probability states, for example, may only have a small impact on the spot price distribution, the implied "cost" of pursuing contract price influence strategies, in terms of sacrificed expected spot profit, is also small. Since these weightings are equal, they cancel (however, note that the assumption of equal weightings, and this cancellation, has an important impact on the interpretation of the second-order condition, discussed in Section 11.3.4).

In the naïve model, an increasing quantity of contracts induced the firm to increase output, since the corresponding decrease in spot price only applied to those units not covered by the fixed price contracts. However, the firm now knows that while a decrease in the spot price does not affect the contracted units directly, it does indirectly by decreasing the mean spot price used in the consumers' DCC, thus decreasing contract revenue. The terms relating to these two effects ($\pm bk_i$) also cancel in the first order condition. Hence we can simplify it to:

$$\begin{aligned} 0 = A - b\left(2g_i^r + g_j^r - L_U\right) - \left(v_i^r e_i + c_i g_i^r\right) \\ +k_i b\lambda\left[2\left(\frac{A}{b} + L_U - \sum_i k_i\right)\left(p^r - E[p]\right) - \frac{1}{4b}\left(3\left(p^r\right)^2 - E\left[p^2\right] - 2p^r E[p]\right)\right] \end{aligned} \tag{11.15}$$

(11.15) shows that, as spot price volatility tends to zero, (i.e., $p^r = E[p]$, for all r) or, equally, consumer risk aversion tends to zero, the first order condition not only reverts to the standard Cournot equation, but also removes any influence of contracts from the optimality condition. While the naïve model assumed that contracts were fixed, regardless of their profitability, we are now assuming the generator anticipates the contract price it will receive. In the absence of spot variability, the contract price will be equal to the equilibrium spot price, hence there exist no (profit making) incentives to sell contracts.

Equation (11.15) also reverts to the zero-contract Cournot solution if the firm were to choose to influence the mean spot price only, i.e., the firm chose not to manipulate volatility to its advantage, but still wanted to influence the expectations of consumers about the mean spot price. In this case, an extra unit of generation will not only reduce the price received for all non-contract units of output (by lowering the spot price), but also reduce the price received for all contract units through its effect on the expected spot price. This is clear in equation (11.14), since all terms in the square brackets relate to the change in variance and covariance, in response to generation, and would be zero for this conjectural variation. Thus the remaining terms describe the zero-contract Cournot solution. In LRE, the contract quantity is equal across all possible states, and these effects counter-balance each other, making the firm indifferent between marginal spot revenue in the current state, and marginal contract revenue in all other states. Hence there is no incentive to increase output above the zero-contract short-run profit maximising output.

Rearranging (11.15) shows that the firm is attempting to balance the change in spot profits with the change in contract profits arising from a change in this state's generation decision, i.e., at optimality,

$$-\left(A-b\left(2g_i^r+g_j^r-L_U\right)-\left(v_i^r e_i+c_i g_i^r\right)\right)=\lambda k_i\left[2b\left(\frac{A}{b}+L_U-\sum_i k_i\right)\left(E[p]-p^r\right)\right.$$
$$\left.+\frac{1}{4}\left(3\left(p^r\right)^2-E\left[p^2\right]-2p^r E[p]\right)\right] \tag{11.16}$$

Since the terms on the left-hand side represent the first order condition for a Cournot firm with no contracts, we know that at the optimal zero-contracting Cournot solution, these terms are zero. At output levels above or below that point, spot profit is being sacrificed, evidenced by these terms being non-zero. It may be helpful to rearrange (11.15) to give an equation for generation:

$$g_i^r=\frac{\left(A-b\left(g_j^r-L_U\right)-v_i^r e_i\right)}{\left(2b+c_i\right)}+\frac{\lambda k_i}{\left(2b+c_i\right)}\left[2b\left(\frac{A}{b}+L_U-\sum_i k_i\right)\left(E[p]-p^r\right)\right.$$
$$\left.+\frac{1}{4}\left(3\left(p^r\right)^2-E\left[p^2\right]-2p^r E[p]\right)\right] \tag{11.15a}$$

Bearing in mind that the generation level will determine the price terms on the right-hand side of (11.15a), this equation shows that, at optimality in each state, the extent to which the firm tends away from the short-run optimal solution (i.e., the first group of terms, describing the zero-contract Cournot solution) is determined by the effect of the generation decision on total contract revenue, through the variance and covariance terms in the contract price. The *direction* of movement away from the zero-contract solution, towards the new solution, is determined by the aggregate sign of the terms inside the square brackets (since we can safely assume that $\lambda k_i/\left(2b+c_i\right)>0$). Two groups of terms are evident inside the square brackets, corresponding to the derivative of variance, and covariance, respectively.

In our evaluations of the individual effects of generation on the variance and covariance, we will consider three "types" of price states: high-price states, where the price is greater than the mean price; low price states, where the price is lower than the mean, and states where the price is equal to the mean.

Perhaps the simplest case is the effect of the variance and covariance terms in a state where the price is equal to the mean. The only consistent generation solution to this scenario is that the firm produces the zero-contract Cournot output. This can be seen by setting $E[\tilde{p}] = p^r$ in equation (11.15a)[73].

The variance term yields an intuitive explanation. If state r is a high price state (i.e., higher than the mean price), corresponding to a low output, and the aggregate level of contracts traded in equilibrium is less than the maximum possible load[74], then the first group of terms is negative, driving the firm's output lower than the traditional Cournot level, thus raising the price. The opposite is true for states in which prices are below the mean - the contract price terms push the generation higher, and thus the price lower.

The change in covariance, with respect to a change in generation, can be found through similar reasoning. If state r is a "high" price state, then we can say that:

$$\begin{array}{ll} & p^r > E[p] \\ \therefore & p^r E[p] < (p^r)^2 \quad \textit{and} \quad (p^r)^2 > E[p^2] \\ \therefore & 3(p^r)^2 - 2p^r E[p] > (p^r)^2 > E[p^2] \\ \therefore & 3(p^r)^2 - 2p^r E[p] - E[p^2] > 0 \end{array}$$

.It is a simple exercise to show that if state r is a low-price state, then:

$$3(p^r)^2 - 2p^r E[p] - E[p^2] < 0$$

[73] Note that this does not imply that there must, in all cost state distributions, be a state where the resulting equilibrium will be the zero-contract Cournot solution. For example, there may be a range of states where none of the actual prices observed correspond to the zero-contract Cournot price, but this reasoning shows that the mean of these prices must be the zero-contract price.

[74] It is reasonable to expect that the consumers will never demand enough contracts to cover the theoretical load they would demand if prices were zero.

Hence, in each type of price state, a change in generation has the opposite effect on variance to that which it has on covariance. To summarise:

- In high price/low generation states, the variance terms lead the firm to decrease generation relative to the zero-contract optimal solution, whereas the covariance term opposes this effect, causing the firm to increase generation, relative to the zero-contract solution

- In low price/high generation states, the variance terms lead the firm to increase generation relative to the zero-contract optimal solution, and the covariance term again opposes this effect, incentivising an increase generation

Stated another way, the variance-related terms in (11.15a) lead the firm to "spread" the prices further than they would be distributed if the firm operated according to the standard zero-contract Cournot policy, whereas the covariance terms are stabilising terms, attempting to push the prices closer together. The net effect of these two effects is difficult to predict analytically (see below). However, this result appears intuitive. As noted in Chapter 9:

> *"...consumers who exhibit optimally-responsive electricity demands are less willing to pay to avoid spot price variance than those who exhibit fixed loads, and thus can't respond." (p 187)*

Since the covariance term is derived directly from the hedging demand of responsive consumers, it should not be surprising that these consumers' contract demand is mitigating the incentives to spread prices, since an increase in covariance reduces the contract price.

If the variance-related terms are larger than the covariance terms, in all states, i.e., if

$$-2\left(\frac{A}{b}+L_U-\sum_i k_i\right)\left(p^r-E[p]\right)>\frac{1}{4}\left(3\left(p^r\right)^2-E\left[p^2\right]-2p^rE[p]\right) \tag{11.17}$$

then the firm will maximise profit by spreading the price distribution wider than the zero-contract Cournot price distribution, although, as long as the covariance terms are non-zero, not as wide as she would achieve without the responsive consumers. The firm would find the solution where the sacrifice in short-run spot profits (the left-hand side of (11.15)) just equalled the increase in contract revenue (the right-hand side of (11.15)). We will call this solution the "destabilising" solution, since the firm is finding increasing the risk for unresponsive customers more profitable than reducing covariance to improve contract prices earned from responsive customers.

If the covariance terms dominates, i.e., the opposite of condition (11.16), the firm's optimal generation policy will be to narrow the distribution, relative to the distribution in the zero-contract Cournot case. We will call this case the "stabilising" solution.

Various rearrangements of (11.16) have been attempted to provide some intuition as to which circumstances will cause each effect to dominate. However, none were appealing, and it would not add anything to the analysis to include them here. We leave this as a numerical question, and, as will be shown in Chapter 12, for reasonable estimates of the parameters, the variance term appears to dominate in all cases.

11.3.3 Cost State Dependencies

Until now, we have developed the first-order conditions assuming that the firms' cost-states are perfectly correlated, i.e., a firm observing cost state r knows, with complete certainty, which cost state its rival is facing. Thus, there is only one combination of first-order conditions that will describe the reaction functions, for a given state, for either firm.

Let us now allow some, or all, of these states to have conditional probabilities less than 1, as described in Chapter 10. In that chapter, we noted that under this assumption, each state observed by firm i has a number of associated potential spot market equilibria.

If the cost states are not perfectly correlated, the derivatives express the average change in variance or covariance, across all possible states being observed by its rival, when the firm is in state r. Hence the derivatives become:

$$\frac{dVar[p]}{dp^r} = \sum_{s \in T_j} 2\theta_i^r \theta_i^{s|r} \left(p^{r,s} - E[p]\right) \tag{11.17}$$

$$\frac{dCov(p,p^2)}{dp^r} = 3\theta_i^r \sum_{s\in T_j} \theta_i^{s|r} \left(p^{r,s}\right)^3 - \theta_i^r E\left[p^2\right] - 2\theta_i^r \sum_{s\in T_j} \theta_i^{s|r} p^{r,s} E\left[p\right] \tag{11.18}$$

These equations can be substituted into the first order condition in the manner described above.

Other than the way in which the terms $E[p]$ and $E[p^2]$ are now calculated, the inclusion of conditional probabilities does not change the form of the firm's first order condition, since the choice of profit maximising response to a given output by other firms is independent of the cost state the rivals are observing.

Having developed the system of first-order conditions, there are two important issues to be addressed. Since we are dealing with much more complex expressions for optimal generation than in the traditional Cournot analysis, we can no longer take it for granted that (a), the solution to (11.15) exists, and defines a profit maximum, i.e., profits are concave in the firm's output level, and (b) the pair of reaction functions implied by solutions to (11.15) intersect to define a unique equilibrium in the spot market. We will now deal with each of these questions in turn.

11.3.4 Profit Concavity

The solution to (11.15) defines a profit maximum for firm i if and only if the second order condition on the solution holds. Taking derivatives with respect to generation yields:

$$0 > -2b - c_i + b\lambda k_i \left[4b\left(\frac{A}{b} + L_U - \left(k_i + k_j\right)\right)\left(1-\theta_i^r\right) - \left(3 - 2\theta_i^r\right)p^r + E\left[p\right]\right] \tag{11.17}$$

Substituting in the inverse demand curve for p^r:

$$p^r = A - b\left(G^r - L_U\right)$$

Simplification allows us to cancel all terms involving A and L_U:

$$0 > -2b - c_i + b^2 \lambda k_i \left(G^r - E[G] - (K - G^r)(2 - 2\theta_i^r) \right) \tag{11.19}$$

Analytical interpretation of condition (11.19), which must hold in all cost states, is difficult, since many of the terms have no clear relationship to one another, and the condition is a function of the decision variables (generation and contracts) themselves, i.e., it changes (linearly) with output and contract level. While a variety of rearrangements of 11.16 were attempted, in order to find an intuitively appealing condition on the parameters, we believed that proceeding in this direction, i.e., an exhaustive analysis of the relationship between parameters and the generation solution, would not add anything to the goal of the thesis. One of the significant intentions of the thesis, outlined in Chapter 3, was to show that stable equilibria exist that support firms behaving in such a way that the distribution of prices is manipulated to increase total profit. Chapter 12 will show that this effect does exist, and, in the context of this thesis, we believe this is sufficient. However, significant analytical investigation of the second-order condition was undertaken, and it is outlined in detail in Appendix B.

It is worth noting the importance that condition 11.16 is met. Recall that a Cournot-Nash equilibrium represents a solution where both firms are maximising profits (i.e., satisfying 11.16), and, for each firm, this solution is stable with respect to their rival's output. Another way of expressing this is to say that the reaction functions, of each firm, intersect. A reaction function describes the profit maximising response of a firm to its rival's output. Normally, even in the situation where reaction functions are non-linear, if we can guarantee that profits are strictly concave, then describing a firm's profit maximising reaction across the entire range of rival output is not problematic (Figure 11.1).

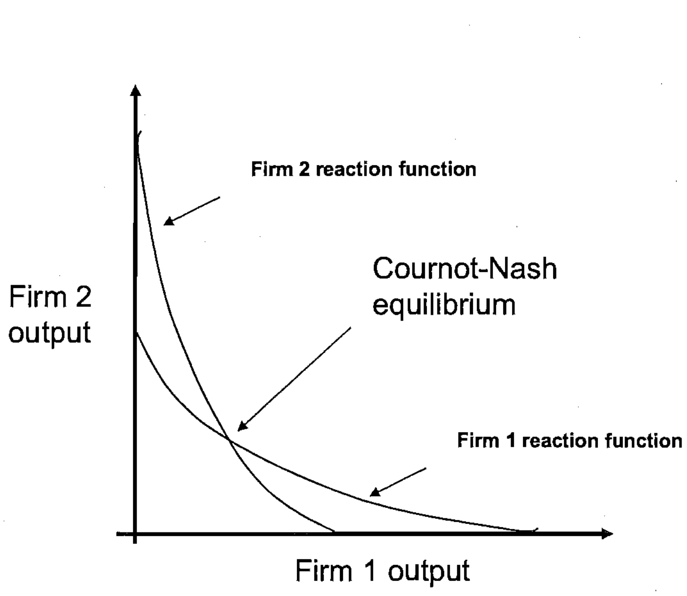


Figure 11.1 Well-behaved reaction functions

However, not only is it not clear whether 11.16 can be satisfied, Appendix B shows that it can be expressed as an upper bound on generation[75]. The implication is that there are levels of generation, for each firm, that are sufficiently high that (interior-point) profit maximising solutions do not exist[76]. Hence reaction functions may not be described over a full range of rival output. These reaction functions may still intersect, thus providing a Cournot-Nash solution (Figure 11.2). However, if the reaction functions do not intersect (Figure 11.3), then there will be no solution to the system of equations.

[75] It is interesting to note that there is no implicit lower bound on generation, in order to guarantee interior point solutions. While we have not been able to prove it conclusively, we believe it is a consequence of prices becoming negative, at high generation levels, with a linear demand curve. Since the first order conditions, described by 11.15, includes terms in both p^r and $\left(p^r\right)^2$, there will be a fundamental change in the way the solution behaves at the point at which prices become negative. While prices are positive, and generation increases, both these terms will decrease. However, when price becomes negative, p^r terms will continue to decrease, while $\left(p^r\right)^2$ terms will begin to increase. This behaviour does not occur for very low levels of generation (and the implied high prices). We would postulate that at high levels of generation, this causes the attractiveness of further price spreading, i.e., increasing generation, and thus the $\left(p^r\right)^2$ terms, to outweigh the cost of doing so, in terms of sacrificed spot profits which is partly described by the p^r terms.

[76] As will be discussed below, this behaviour is most likely because further price spreading becomes increasingly attractive, as generation increases and prices decrease. Hence profit is likely to be continuing to increase past this upper bound on generation, but at an increasing rate. Hence condition 11.16 is a condition on an interior point profit maximising solution. It is worth noting that with well posed cost functions, and/or bounds on generation, this problem could be avoided.

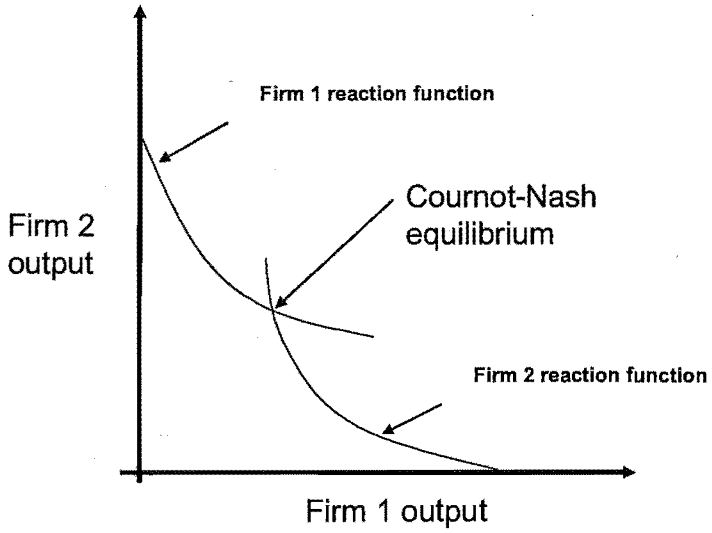


Figure 11.2 Cournot-Nash equilibrium with ill-behaved reaction functions

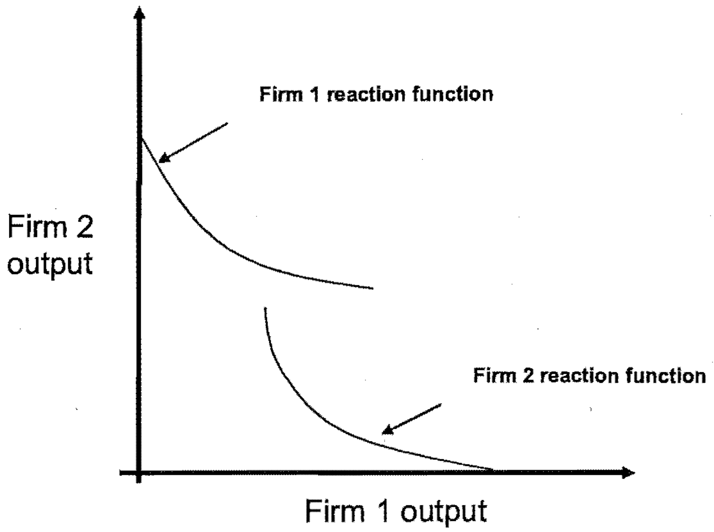


Figure 11.3 No Cournot-Nash equilibrium with ill-behaved reaction functions

Rather than attempt to provide numerical estimates on the relative magnitudes of the critical terms in 11.16, we instead will present two highly simplified numerical examples, one in which an equilibrium is observed, and one where it isn't.

11.4 Two-state example

Consider a situation where each firm faces two possible cost states. Both firms face the same cost state simultaneously, and each state has the same effect on each firm's cost intercept, thus we have four generation first-order conditions, which must be solved simultaneously. The parameters are shown in Table 11.1.

	e_i	c_i
Firm 1	\$20/MWh	\$0.06/MWh
Firm 2	\$40/MWh	\$0.04/MWh

Table 11.1 Example Firm Cost Parameters

In the low cost state (state 1), both firms' cost intercept is decreased by 10%, and in the high cost state (state 2), the intercepts are increased by 10%, i.e., v_i is 0.9 and 1.1, for $i = 1, 2$. The spot demand curve has intercept A = \$120/MWh, and slope b = \$0.02/MWh.

It can be shown that the zero-contract Cournot solution, in each state, would lead to prices of \$94/Mwh and \$93/MWh for the high and low cost states, respectively. However, through first order condition 11.15, the risk-manipulating incentives acting on each firm cause them to shift away from this solution. For this set of parameters, the variance effect dominates the covariance effect, and the firms settle on a destabilising policy where they produce less in the high-cost state, and more in the low-cost state, and spread the prices to be \$97/MWh and \$90/MWh, respectively.

While the purpose of this example is to illustrate the behaviour of the generation first-order conditions, it is worth noting the effect of this strategy in the contract market. The firms succeed in increasing the variance of the spot price from \$1/MWh^2 to \$20/ MWh^2, and while the covariance is similarly increased, which has a negative effect on contract prices (in isolation), the net effect is an increase in contract prices by \$0.50/MWh, and thus contract profit.

Let us now consider a situation where the firms have very flat marginal cost curves, which are more dramatically affected by the cost state. Table 11.2 summarises the firms' cost functions.

	e_i	c_i
Firm 1	\$20/MWh	\$0.01/MWh
Firm 2	\$40/MWh	\$0.01/MWh

Table 11.2 Example Flat Marginal Cost Parameters

In this example, firms will experience an increase/decrease in costs, relative to the figures given in the table, of 40%, in the high and low cost states, respectively[77].

The zero-contract Cournot solution, and the net effect of the risk manipulating terms at that solution (i.e., the value of the terms inside the square brackets in 11.15), are shown in Table 11.3.

	Zero-Contract Generation. Low Cost	Zero-Contract Generation. High Cost	Risk-Manipulating Effect, Low Cost	Risk-Manipulating Effect, High Cost
Firm 1	1800MW	1724MW	274MW	-262MW
Firm 2	1400MW	790MW	274MW	-262MW

Table 11.3 Equation 11.15 terms

The price in the low- and high-cost states is \$66/MWh and \$79/MWh respectively. Again, the variance effect dominates, and each firm is encouraged to spread the price distribution wider than the zero-contract Cournot solution.

[77] These parameters, while unrealistic in a setting where capacity constraints aren't considered (particularly the identical slopes of the cost curves), have been deliberately chosen to encourage generation solutions, in the low-cost state, to be high, in order to illustrate how the behaviour of the first order condition changes at very low prices.

In order to observe the path that the four generation first-order conditions take, in the search for a simultaneous solution, we will perform a single iteration of a Newton-style algorithm, by adding the risk-manipulation terms to the zero-contract Cournot output, to obtain a new solution. The results are displayed in Table 11.4.

	New Generation. Low Cost	New Generation. High Cost	Risk-Manipulating Effect, Low Cost	Risk-Manipulating Effect, High Cost
Firm 1	2074MW	1462MW	727MW	-652MW
Firm 2	1674MW	528MW	727MW	-652MW

Table 11.4 Second iteration, Equation 11.15

While the price in the high-cost, low generation state has been raised to \$90/MWh, the price in the high-generation, low cost state is now \$55. The variance of the spot price has increased from \$94/ MWh^2, in the initial solution, to \$619/ MWh^2. Furthermore, as shown in the table, the magnitude of the terms encouraging each firm away from the zero-contract Cournot solution has increased markedly. Clearly, further destabilisation is increasingly attractive. Subsequent iterations show that these terms magnify at an accelerating rate, and it is obvious that no equilibrium can be found.

It is clear from the example that, in the absence of sensible bounds on generation (either implied by the steepness of the marginal cost function, or physical bounds on the generation variable), finding simultaneous solutions to the problem presented here can be problematic. The parameters chosen in the second example, were relatively extreme, however, implying almost constant, and low, marginal costs, with no upper bound on generation.

It is also worth noting that the likelihood of an equilibrium being found in the spot market is improved if the firms are not guaranteed to simultaneously face identical costs states (implied by a perfectly correlated cost distribution), since there is, potentially, less chance that both firms will simultaneously face low costs, and therefore both be attempting to increase generation. While the firm in the low-cost state will face the same reaction function complications described above, its rival may be able to form a reaction function over a wider range of generation levels, in those states which do not provide strong incentives to increase generation. Nevertheless, as long as there is a positive joint probability that both firms will face the low-cost state(s) simultaneously, the difficulties with defining reaction functions, and finding equilibria, remain.

While the parameters chosen initially resulted in a stable equilibrium, this was a highly stylised example, and since numerous parameters are present in the second-order condition 11.16, interacting in a complicated manner, we have obtained little information as to how to guarantee stable and sensible equilibria. We will now consider ways in which the above approach to solving the generation first order conditions described by 11.15, and the contract first order conditions 11.3, can be changed to assist finding solutions.

11.5 Implications for Solving the LRE

The discussion above has outlined the incentives acting on a firm that wishes to influence contract prices by altering its spot market strategy from the traditional short-run profit maximising one. We have shown that under the assumptions of the LRE framework, first-order incentives to manipulate the risk faced by both responsive and unresponsive consumers push generation, in each state, away from the traditional zero-contract output.

However, we also established that profit, for a destabilising firm, is not always concave, and hence finding solutions to the first order conditions is potentially problematic. Essentially, the second order condition represents a limit on the firm's ability to increase generation above the mean. Appendix B shows that how restrictive this condition is depends on the combination of a number of parameters, in particular, the contract level, consumer risk aversion, and the probability of the state occurring.

We will now consider how we might address the instability of the solution, particularly at high levels of generation (and thus low cost states). First, we will introduce the notion of downside risk, and how we will implement it in the model. Second, Section 11.5.2 will propose that we also consider the possibility that only one of the firms destabilises.

Section 11.5.3 will outline other possible resolutions to the issue of profit convexity which will not be considered here.

11.5.1 Downside Risk

We argued above that the most likely states to be restricted by the second-order condition are those for which we expect generation to be higher than the mean generation, i.e., low cost states. It is in these states that the firm faces first-order incentives to push generation higher than the short-run profit maximising level, and thus closer to the bound implied by. One could reasonably question the conjectural basis of a firm attempting to create risk for consumers by driving prices down. As discussed in earlier chapters, and evidenced by actual market behaviour, the states most likely to be of concern to consumers are those in which prices are high, driven by high generator costs and thus low generation. Additionally, hydro generators are more likely to feel comfortable with strategies that increase storage, rather than empty it.

The fact that our model creates incentives for generators to spread prices in **all** states is a direct result of using statistical variance as a measure of risk, which implies that the frequency of low-price states contributes equally to the overall risk assessment as the frequency of high-price states. Thus we have a sound rationale for ignoring destabilisation in the problematic (i.e., low cost) states altogether, i.e., by incorporating downside risk into the model. We will now propose a measure of downside risk that allows us to retain the exact same expressions developed above, as optimality conditions, for the problem facing the generators.

It initially seems appropriate to reformulate the model based on a measure of downside risk, such as semi-variance. However, these discontinuous measures of risk make the development of the responsive consumers' DCC complicated, particularly the calculation of covariance, which was critical to those consumers' ability to vary load in response to the spot price. We will adjust the model to include a much simpler representation of the consumers' assessment of risk, and thus their demand for contracts. While the measure of risk we propose has no established theoretical basis, established measures of risk are at best a hypothetical approximation of a very complex psychological process. In any case, any measure of variability that only considers outcomes which are "bad" for the decision maker is an improvement on the use of statistical variance.

Consumers choose which states they consider to be a risk, and we suggest this will be done by deeming the states in which prices are above the median to be considered "risky". Since the responsive consumers' profits, and the unresponsive consumers' costs, are, respectively, decreasing and increasing in price, this also corresponds to the consumers choosing states in which profits (costs) are below (above) the median. States with prices below the median, while contributing to the variance and covariance of the overall distribution, will not be considered for this measure of "pseudo-risk". Consequently, generators will choose only to spread the prices in those states which will increase risk, i.e., those states with costs above the median. We will continue to use the traditional expressions for variance and covariance, except that their calculation will only include those observations from the states which are considered risky, although the deviations of these individual observations will be measured from the overall mean. Mathematically, we will define a set D, that contains the states that are considered to be risky, i.e., $r \in D$ if and only if $p^r > med(P)$, where P is the full vector of prices observed in the market, and *med(P)* describes the median of this distribution. Then our measures of variance and covariance become:

$$Var[\tilde{p}] = \sum_{t \in D} \theta_t \left(p_t - E[P]\right)^2 \tag{11.20}$$

$$Cov\left[\tilde{p}, \tilde{p}^2\right] = \sum_{t \in D} \theta^t \left(p^t - E[P]\right)\left(\left(p^t\right)^2 - E\left[P^2\right]\right) \tag{11.21}$$

Since they will only be relevant to states in which price spreading takes place, the derivatives of these measures, developed in Section 11.3.2, remain the same.

From a computational perspective, we need to ensure consistency between the consumers' choice of risky states, and the equilibrium prices that result from the solution of the model. However for the correlated cost situation, prices will be ordered, and distributed, according to the cost distribution. Hence the states with prices above the median price will correspond to those states with marginal cost intercepts above the median intercept[78]. The makes the assignment of risky states relatively straightforward.

However, whether a firm should employ price spreading in a given risky state is not so obvious. The first order incentives above show that generation firms will push prices further away from the mean. If we apply these incentives to all high-cost states, we would expect that the median, and thus the set of risky states, would be unchanged under price spreading. However, since the variance and covariance measure the deviation of a given price from the *mean*, the direction of spreading is still determined by the price relative to the mean. It is conceivable that, if the highest cost states result in very high prices, the mean will increase enough to cause a moderate-cost risky state to have a price below the mean. Thus a firm will face incentives to lower this price, possibly affecting the assignment of risky states.

Both the assignment of risky states and destabilisation incentives are more likely to result in inconsistencies when independent cost distributions are introduced, as a firm being in a high-cost state does not then imply *a priori* that the state is "risky", and thus destabilisation is profitable.

To account for these computational aspects, an algorithm is introduced, in Chapter 12, which seeks consistency between the assignment, and the realisation, of risky states.

In the states not considered risky, the firms will not attempt to spread the prices. However, this does not imply that the firms behave "naively", as described in Chapter 10. Rather, the firms will produce the zero-contract output, recognising the effect that spot prices in a given state have on future contract prices via the mean spot price.

11.5.2 Different Firm Strategies

We will also consider the possibility that both non-destabilising and destabilising firms may simultaneously exist in the market. While not guaranteeing it, this increases the

[78] Equally, we could set an initial assignment by calculating the zero-contracting Cournot prices, since the first order conditions showed that with symmetric spreading, the destabilised equilibrium price will be further away from the mean than the zero-contract price.

chance of an equilibrium in the spot market, since removing the terms arising from destabilising incentives results in well-behaved, linear and downward sloping reaction functions, according to:

$$\hat{g}_i^t = \frac{A - b\left(g_{-i}^t - L_U\right) - v_i^t e_i}{\left(2b + c_i\right)} \tag{11.22}$$

i.e., the traditional zero-contract Cournot output. Note that just because a firm chooses not to destabilise, it is not necessarily "naïve" as defined in Chapter 10[79]. Non-destabilisation here simply implies that while the firm is fully aware of the impact spot outcomes have on contract prices, it chooses not to spread prices. The destabilising firms face the same problem as discussed above: that a reaction function may not be defined over the full range of competitor output, and may not exist at all. However, assuming that the latter case does not apply, (11.22) is defined over the full range of output, and may intersect the destabilising firm's reaction function, especially given our use of pseudo-variance.

11.5.3 Constraints on Destabilising Behaviour

Even though the above analysis has shown that a firm may face incentives to infinitely spread the price, in reality, there must be a limit to this type of behaviour.

In reality, capacity constraints, bounds on prices and/or the threat of regulation or entry would limit the extent to which prices can be increased (or decreased) from their short-run profit maximising level. Firms could have an upper bound on the price volatility they were able to create, either as an internally-set constraint to avoid regulation, or by regulation itself. Such factors would lead to constrained profit maximising solutions to (11.15) for each firm, even though profits were not concave. This would, however, lead to a discontinuity in the reaction functions, at the point at which the constraint becomes binding, although this, in itself, does not preclude equilibria from existing. While the

[79] Although this could be another modelling alternative, with the same increase in the chance of an equilibrium as for equation (11.22)

model proposed above does allow these sorts of extensions, we consider it an area for future development of the model, and it will not be investigated further here.

11.6 Conclusions

We have extended the LRE model presented in Chapter 10 to allow the supply firms to deliberately influence the DCC. In particular, the firms adjust their spot market behaviour so that the effect each generation decision on the consumers' observed distribution of spot prices is optimised in a total (i.e., contract and spot) expected profit maximising way.

However, the above analysis shows that the spot market problem faced by a firm who wishes to influence the contract market in this way does not necessarily yield well-behaved solutions. Neither does it present us with any clear analytical bounds on the parameters, under which solutions will be well behaved and equilibria found. We believe that this, in itself, is an important conclusion, given the aims of this thesis – that the problem of firms influencing both contract and spot markets, simultaneously, is not an easy one to solve, and does not necessarily yield stable equilibria. It should not, therefore, be surprising to find that numerical models of the problem, especially those developed more elaborately that the one presented here, do not find solutions in certain situations

Notwithstanding these issues, a number of important analytical results were developed. Firstly, given the effect of an individual state's spot price on the mean spot price, a positive level of contracts no longer provides explicit incentives for the firm to adopt a more competitive strategy on the spot market, since any marginal revenue losses avoided in the spot market, from having output sold on contract, are realised in the contract market. Hence the firm does not attempt to influence the mean spot price directly, rather, it anchors its output to the standard zero-contract output, thus implicitly raising the mean spot price.

Secondly, the first order conditions developed showed that a firm that considers the impact of its spot actions on the contract prices faces an incentive, from the unresponsive consumers' contract demand, to "spread" prices, around the mean, wider than the zero-contract optimal output. On the other hand, the responsive consumers' ability to manage risk by altering their level of activity resulted in a counteracting incentive on the generators, leading them to stabilise prices, relative to the short-run optimum. The net effect of these two effects remains a numerical question, and will determine whether the spot price distribution becomes more or less risky for consumers, than it would if the generation firms chose to act according to Cournot conjectures, in the absence of contracts. While the firm is no longer acting in short-run spot market optimality, these losses are compensated for with increased contract prices and contract revenue. Contract prices are improved by reducing the covariance of the spot price, with its square, or increasing the variance of the spot price, but, as highlighted above, a firm cannot achieve both objectives simultaneously.

However, the profits from this manipulative behaviour are not strictly concave in the firm's own output. Concavity would imply that the trade-off between spot losses and contract gains would equilibrate at some level of generation, providing a stable solution. If the net incentive is to "spread" prices, and is attractive enough, profits may become convex at high levels of generation, implying that a profit-maximising firm faces incentives to infinitely spread the price. The concavity condition can be rearranged to form an upper bound on generation, or the spread of generation from the mean, in a particular state. While the first order condition may encourage firms to increase output above the mean, in low-cost states, this, in fact, drives the problem behaviour further away from concavity, and thus away from a *solution* to the first-order condition. An attempt at quantifying the behaviour of the first- and second-order conditions, and the resulting impact on the Cournot game, is presented in Appendix B, but did not yield results that were satisfactory enough to warrant reporting as part of this thesis. Here, we leave the question of how to describe the region of the solution space that will yield concave profit to future work, and will rely on numerical results to illustrate that equilibria can be found.

In order to address the concavity difficulties, a number of model enhancements were suggested (e.g., the inclusion of capacity constraints), but it was decided that the most convenient, and simple, adjustment to the model would be to include a measure of downside risk, instead of the full statistical variance. Downside risk is more intuitively appealing, since only high price states are likely to be considered risky by the consumers. This would reduce the number of states that destabilising is appropriate in to those with high generation costs. It is in these states that the generators' output is lowest, and thus least likely to yield a violation of the concavity condition.

Thus a concept of "pseudo-variance" was introduced, where only states in which prices were above the median are considered to be risky by the consumers. Only in these states will generators attempt destabilisation; else, they will act according to the zero-contract Cournot equation. We acknowledge that the particular measure of downside risk, presented here, is not an established measure. However, this measure is surely no less plausible than the assumption that consumers' risk attitudes are driven by the more traditional, symmetric measure of risk. In any case, we do not believe that it detracts from the aims of this thesis, nor the significance of the results presented in the next chapter.

Unfortunately, while we believe these enhancements will increase the likelihood of equilibria, we still cannot guarantee it as we were not able to develop any explicit conditions on the parameters involved, in determining when stable solutions can be found. Thus it remains to explore the solution behaviour numerically, under each of the scenarios outlined above, which will be addressed in the next chapter.

12

RESULTS

12.1 Introduction

The previous two chapters have proposed two models of firm behaviour in a market that has reached a state of long-run equilibrium. The first, which we referred to as the "naïve" model, found contract and spot market equilibria where firms ignored the effect the range of market outcomes has on contract prices, while the second suggested that firms will deliberately influence the spot price distribution so as to advantage their contract profits. It was shown that firms behaving under the latter scenario faced competing first-order incentives: one to spread prices, one to stabilise prices.

This chapter intends to numerically illustrate the effects of these incentives. We will benchmark the results against the market outcomes achieved in both a perfectly competitive and Cournot situation, where no contract market is present.

While we provide numerical solutions to the LRE systems, we do not intend to provide a description of the actual equilibria that would be observed in any given market. As discussed in Chapter 10, the LRE framework itself is only a hypothetical "average" state of the market in the long-run, and does not describe actual market behaviour in any particular state. Additionally, the parameters we use in the model have not been benchmarked with empirical data from any particular market, although they will be chosen to reflect a reasonable range of industry structures.

Instead, we intend to lend numerical support to the insights that have been developed in the previous chapters, mainly that firms who are aware of the impact of their spot strategy on consumers' risk perceptions may attempt to "destabilise" the spot market, i.e., the incentives provided through load-responsive demand for contracts is weak relative to unresponsive contract demand. We will show that the extent to which this behaviour occurs, and is profitable, depends on the values of key underlying parameters, in particular the variability in the underlying cost distribution and the degree of consumer risk aversion.

Section 12.2 will outline the underlying parameters that will be used in the models that follow, and the type of markets that they reflect. Section 12.3 will present results for both the perfectly correlated cost-state, and the independent cost state case. Sections 12.4 and 12.5 will attempt to quantify the relationship between cost variability, and consumer risk aversion, respectively, on the degree of equilibrium price-spreading. Section 12.6 provides concluding comments.

12.2 Parameters

12.2.1 Generators

Generators' marginal cost functions are assumed to be linear with a positive intercept. In the base case (ignoring uncertainty in the intercept), Firm 1 has a marginal cost intercept, e_1, of \$20/MWh and a slope, c_1, of \$0.06/MWh. For Firm 2, e_2 = \$40/MWh and c_2 = \$0.04/MWh. While it is hard to give an intuitive interpretation of these values in terms of a marginal water value curve for a hydro firm (since water values are, in reality, the result of a reservoir optimisation), Firm 1's parameter values could approximate the supply curve of a thermal firm with cheap baseload plant, and expensive peaking plant, and Firm 2's a mid-merit order range of plant, with slightly cheaper peaking plant (Figure 12.1).

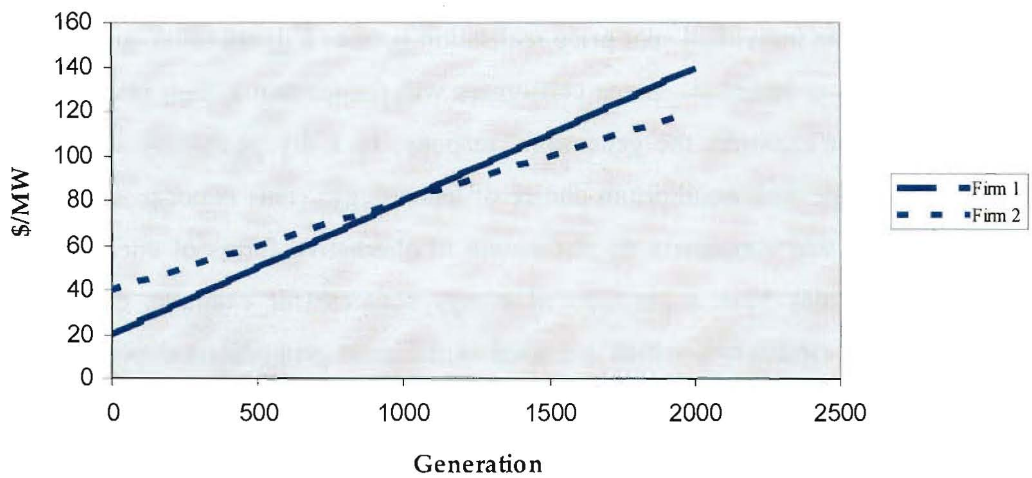


Figure 12.1 Marginal cost functions

The impact of variation in the marginal cost functions will be discussed in the sections detailing the cost state distributions.

12.2.2 Consumers

Recall that, in order to reflect the true volatility in prices, rather than predictable trends in prices, the model should be solved for each subperiod of the load duration curve. The results presented here will be for an individual subperiod, and while it is noted that a multi-subperiod analysis could be performed, it would not add anything to the intentions of this chapter.

Responsive consumers submit a linear aggregate demand curve. Combined with average unresponsive load of 750MW, the parameters A and b were chosen to given an elasticity of -0.25[80], at an approximate perfect competition price of \$70/MWh. This elasticity is chosen to reflect changes in firms' electricity demand, in response to the price, over the medium to long term. For periods under a day or a week, we would expect the elasticity to be much lower (almost inelastic), but the model is intended to reflect an average generation level over a contract period (usually about 12 months), rather than changes within it. Under the LRE framework, the elasticity should mirror the consumers' knowledge that an individual spot price realisation is from a distribution, and should not be responded to in isolation. Since consumers will realise that a high price in a given year may be, for example, the generators' response to a dry year, they should not be inclined to change their equilibrium choice of technology. This is not to say, however, that we assume that consumers do not switch to alternative forms of energy, since an individual firm may have a portfolio of energy sources (for example, co-generation, natural gas and electricity), which are used in different proportions depending on the prevailing (mean) spot price in a year. The elasticity used here is assumed to reflect this, but not any significant changes to this portfolio: we assume the firm has established its portfolio of technologies in a long-run equilibrium.

As modelled in Chapter 9, consumers display mean-variance preferences, and select a contract quantity which maximises expected utility. Chapter 11 motivated the use of what we termed "pseudo-variance", where consumers only consider states in which the price is "bad" (i.e., high), to be risky. While this is similar to measures of semi-variance, we have chosen risky states to be those with prices above the median. Since we make no specific assumption about the shape of the distribution of the cost states (and thus their profits or costs), and we retain the traditional quadratic form of statistical variance (in risky states), we can assume that this reflects an underlying quadratic utility function[81]. Responsive profit, within a given hour, is approximately 1×10^5, in the competitive case, hence the coefficient of relative risk aversion, λ, is 0.0002, in order to reflect an "average" responsive relative risk aversion of between 1 and 2. Unresponsive consumers will be assumed to exhibit the same risk aversion[82].

Since the slope of the contract demand curve is the product of the coefficient of relative risk aversion and variance, the effective elasticity of the contract demand curve will vary

---

[80] This is supported by an empirical study by Dahl (1996) on elasticities of electricity load for a range of consumption types in the U.S. Few studies exist in New Zealand that could provide more reliable estimates.

[81] If profits or costs are normally distributed, a mean-variance utility function reflects an underlying exponential utility function, with constant relative risk aversion.

[82] Since unresponsive retailers' average electricity costs, in the model, are around 1×10^6, this value of λ gives rise to a similar value of relative risk aversion in profit if we assume that retailers use a markup of approximately of 10%. These "normal" values of relative risk aversion are taken from an empirical evaluation of financial managers' risk aversion by Mehra and Prescott (1985)

as generators create more or less variance in the market. For the range of variances given in the results that follow, the elasticities range from -1.4 to -0.5, as variance increases, indicating a higher degree of elasticity in the contract market than in the spot market. This seems reasonable, since the contract decision reflects a tradeoff between exposure to the spot market, and the certainty of fixed price contracts, over a long time frame.

12.3 Results

Results for the LRE system were generated for the following scenarios:

- Perfect Competition, with no contract market, denoted "PC" on the figures.

- Cournot conjectures, with no contract market, denoted "Zero Contract". This also represents the spot equilibria we would observe if both firms were aware of the effect their generation strategy had on contract prices, but chose not to destabilise (i.e., they optimise the effect on the mean, but not the spread of the spot price distribution). Chapter 11 showed, that under this scenario, firms revert to generating as though they had no contracts, a result reported in Green (1993).

- Cournot conjectures with a contract market, and firms acting "naively", i.e., setting spot output assuming that the level, and price, of contracts is fixed, and ignoring the effect spot outcomes have on contract prices through the pseudo-variance of the spot price. This strategy will be labelled "Naïve" on the figures.

- Destabilising Cournot, with both firms destabilising in the high-price states, denoted "Dest, 2F". Given that prices are most likely to be above the mean in high-cost states (although, this is not always true, see Section 12.5), generators choose to destabilise in any state in which the cost-effect, $\tilde{v}$, is greater than 1. In the other states, generators set output as for zero-contract Cournot competitors. This reflects the fact that, while they choose not to destabilise, they are still aware of the effect spot outcomes have on the mean spot price, and thus contract prices.

- Destabilising Cournot, with Firm 1 destabilising in the high cost-states, and Firm 2 choosing not to destabilise at all ("Dest, 1F"). However, Firm 2 is not "naïve", because it is aware of the effect contracts have on the mean spot price, and thus sets output as though it were a zero-contracted Cournot competitor. As will be seen later, the independent cost scenario results in only one firm destabilising, in some states, by default. Thus we will only investigate this as an individual policy for the correlated cost case.

Each of these scenarios corresponds to a particular system of equations, with the destabilising cases represented by two types of equation sets, depending on whether they are destabilising in that particular state or not. The systems were formulated in GAMS (Brooke, Kendrick and Meeraus (1992)), and solved using the CONOPT (Drud (1992)) non-linear solver.

12.3.1 Correlated Cost States.

First, let us consider the situation where the firms face a perfectly correlated probability distribution for $\tilde{v}$. We shall let $\tilde{v}$ have 8 states. We model this as a symmetrical, and approximately normal, discrete distribution, as illustrated in Figure 12.2. To simplify this example, we assume that the effect on each firm's costs is also identical.

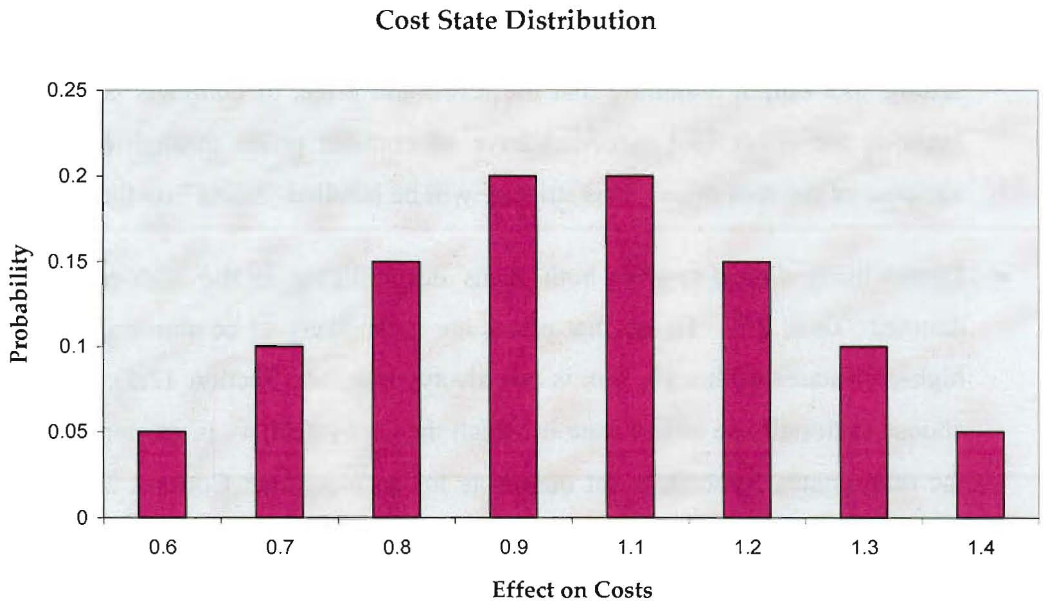


Figure 12.2 Distribution of cost effects

While both firms face the same scaling effect of the cost state on the intercept, note that the differences in base cost assumptions mean that the resulting absolute variability in marginal cost differs between firms. The above distribution results in a standard deviation of marginal cost, at an output of 1000MW, of \$4.50/MWh for Firm 1 and \$8.90 for Firm 2.

Spot Market Equilibria

The equilibrium spot price in each of the eight states is graphed in Figure 12.3. As expected, the competitive price is the lowest, and the effect of a positive level of contracts (see below) results in naïve pricing between the competitive and zero-contract level. The price-spreading effect is clear in both destabilisation cases, with prices being pushed highest when both firms attempt to spread.

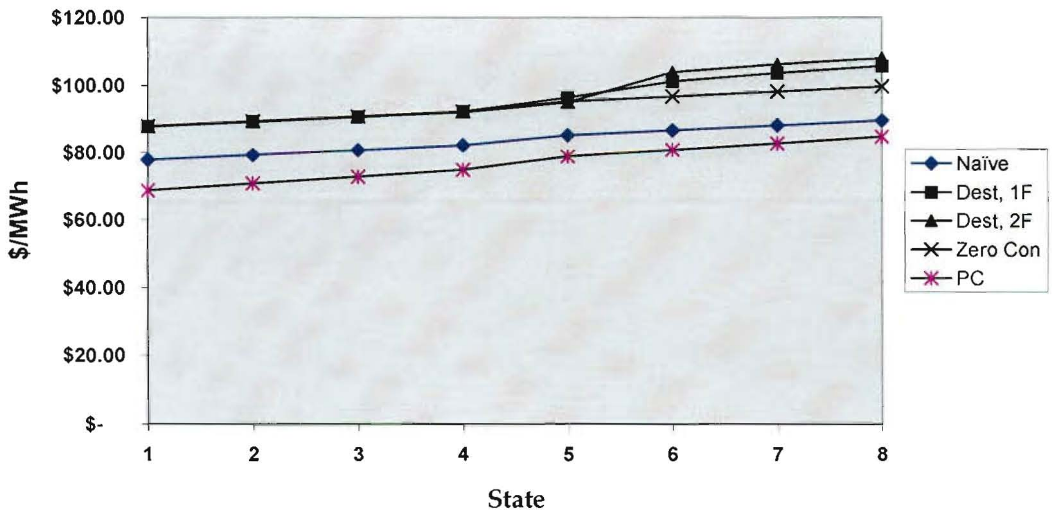


Figure 12.3 Spot price, correlated cost scenario

The resulting mean and variance of price is given in Table 12.1. Recall that, in the risk-manipulating model, firms follow the zero-contract Cournot strategy in the non-risk states, but face incentives to vary from this output in the risky states. Figure 12.3 shows that the firm is clearly spreading prices away from the mean in the latter states, implying that the destabilising incentives are dominating. Hence the average spot price under the one-firm and both-firm destabilising models are higher than under the other scenarios.

	PC	Zero Contract Cournot	Naive Cournot	Firm 1 Destabilising	Both Firms destabilising
Mean Spot Price	$ 76.91	$ 93.91	$ 83.75	$ 95.64	$ 96.04
Quasi-Variance Spot Price	$ 10.06	$ 5.51	$ 5.51	$ 16.78	$ 27.60

Table 12.1

In Chapter 8, we argued that dominant firm behaviour, in the absence of contracts, serves to stabilise prices, which is supported by the results in the table above. However, the results also illustrate how dominant firms with the desire to increase contract profits through spot price volatility may make spot market exposure more risky for consumers than in the competitive case. Whether consumers' overall risk position is increased by generators' destabilisation depends on their optimal contract strategy (see below).

The spot price is, of course, driven by the underlying variation in the generators' output. Figure 12.4 and Figure 12.5 illustrate this.

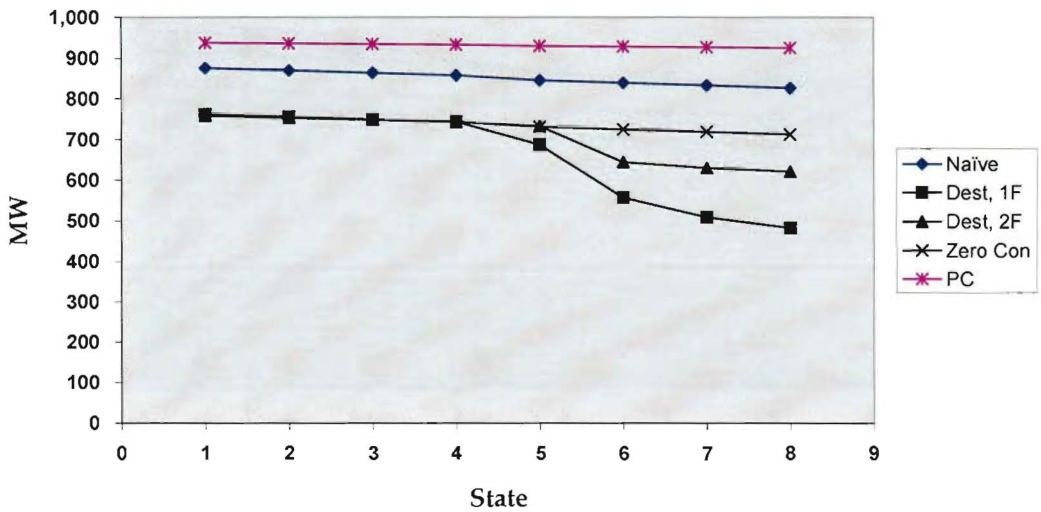


Figure 12.4 Generation, Firm 1

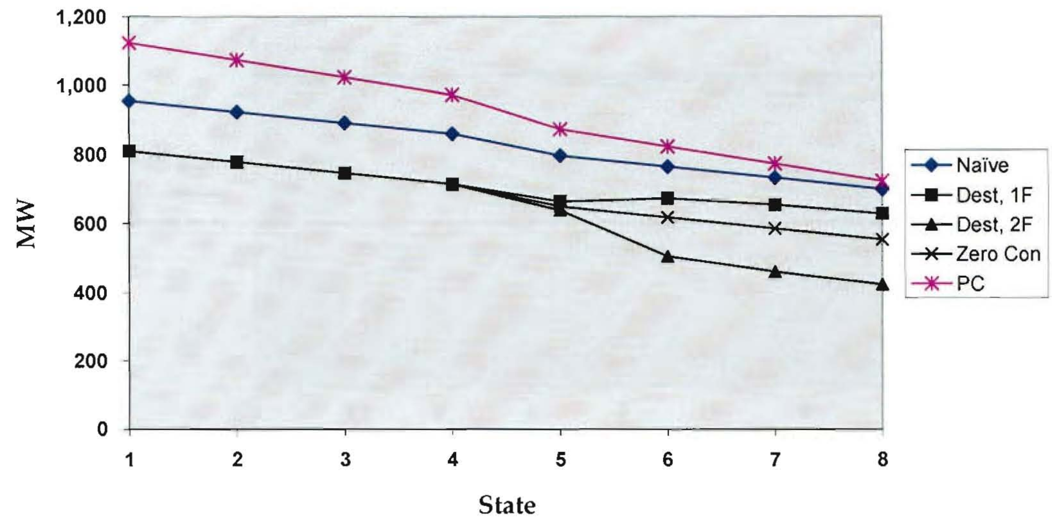


Figure 12.5 Generation, Firm 2

These figures reveal an interesting strategic interaction between the firms, in the case where only Firm 1 spreads prices. Figure 12.4 shows that Firm 1 pursues destabilisation to a far greater extent in states 6, 7 and 8, than it does when both firms destabilise. Figure 12.5 reveals why this is so. Since Firm 2 is not destabilising, it responds to Firm 1's output choices according to the traditional, downward sloping Cournot reaction function. The greater the reduction in Firm 1 output, the higher the profitable generation for Firm 2 is. From Firm 1's perspective, not only does it have to spread generation, it must also counteract the stabilising response of Firm 2, in order to achieve a certain price spread, which comes at a much greater cost. However, when both firms destabilise, their respective efforts at spreading prices support each other, and each, individually, can spread generation to a lesser degree, to achieve the same effect.

However, this is not true of state 5. In this state, Firm 1's spread of generation is considerably lower, when it destabilises by itself, than when it is supported by Firm 2. Figure 12.6 reveals the cause of this.

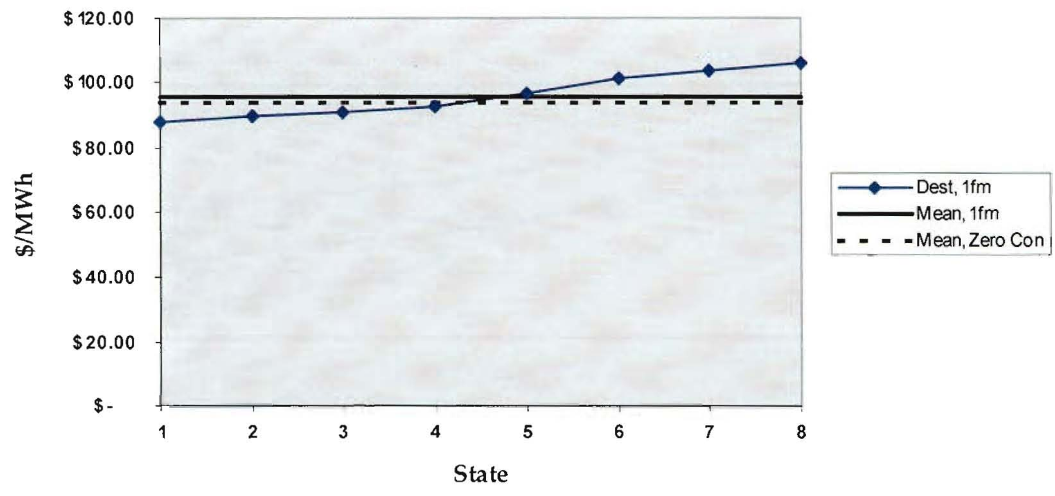


Figure 12.6 Spot price and mean spot price, Firm 1 destabilises

Through the generation spreading strategy of Firm 1 in high-cost states, the mean spot price has been raised above the level observed when neither firm destabilises. Given the stabilising behaviour of Firm 2 described above, Firm 1 does not find it profitable to spread generation in state 5, or, equivalently, the extent to which it would have to force output away from the mean, in order to make a noticeable effect on the spot price distribution, would not be profitable. However, when Firm 2 destabilises as well, spreading is profitable for Firm 1 in state 5, since it is supported by its rival, thus together they achieve a greater effect on the price variance.

Contract Market Equilibria

Contract market results are shown in Table 12.2. The first-order conditions, developed in Chapter 10, showed that both generators would sign the same quantity of contracts. Since both generators face the same contract demand curve, and the implied "cost" of contracting (i.e., the spot price not received on the marginal contract unit) is also identical between firms, the marginal contract revenue curves will be identical for both firms. However, different levels of spot volatility give rise to different DCC's, and Table 12.1 shows that the quantity of contracts sold in equilibrium decreases as the spot volatility increases. This initially seems counter intuitive. But it is the pseudo-variance that forms the slope, and thus the elasticity of the contract demand curve. As in a spot market, a more inelastic demand curve leads to decreases in profit maximising output. Here, as pseudo-variance increases, the contract price elasticity decreases, and the profit maximising quantity of contracts traded in equilibrium decreases.

	Naïve Cournot	Firm 1 Destabilising	Both Firms destabilising
Contracts, Firm 1 (MW)	554	441	418
Contracts, Firm 2	554	441	418
Contracts, Responsive	609	338	281
Contracts , Unresponsive	472	529	541
Contract Price	$ 84.55	$ 97.57	$ 103.52
Risk Premium	$ 0.80	$ 1.93	$ 4.25

Table 12.2: Contract market results

Given their respective abilities to respond to spot price volatility, responsive consumers purchase fewer contracts, to optimally hedge risk, than unresponsive consumers. The risk premiums that result from the spot price pseudo-variance (and covariance with responsive consumer profit) are not significant in the naïve case, approximately 1% of the total contract price. Clearly, the relatively low variability in spot prices when firms behave in ignorance of spot-contract dynamics leads to an elastic demand for contracts, with little scope for pushing the contract price above the mean spot price. Hence generators sell a moderate amount of contracts relatively cheaply.

However, as spot price variability increases, contract demand becomes more inelastic. Hence generators find it more profitable to restrict the quantity of contracts they make available for sale, and consumers' utility-maximising response leads them to offer much higher risk premiums. In equilibrium, generators are extracting a 2% premium in the one-firm destabilising case, and twice that when both firms destabilise.

Whether consumers are more or less contracted, at optimality, as the generators behave more aggressively, depends on their ability to respond by changing their electricity purchases. Since the destabilising behaviour leads to higher spot prices, responsive consumers will purchase less electricity as they reduce their profit maximising output. Unresponsive consumers do not, however. Given the consumers' utility maximising choice of contract level, Figure 12.7 illustrates the proportion of their average electricity load that is covered by long-term contracts.

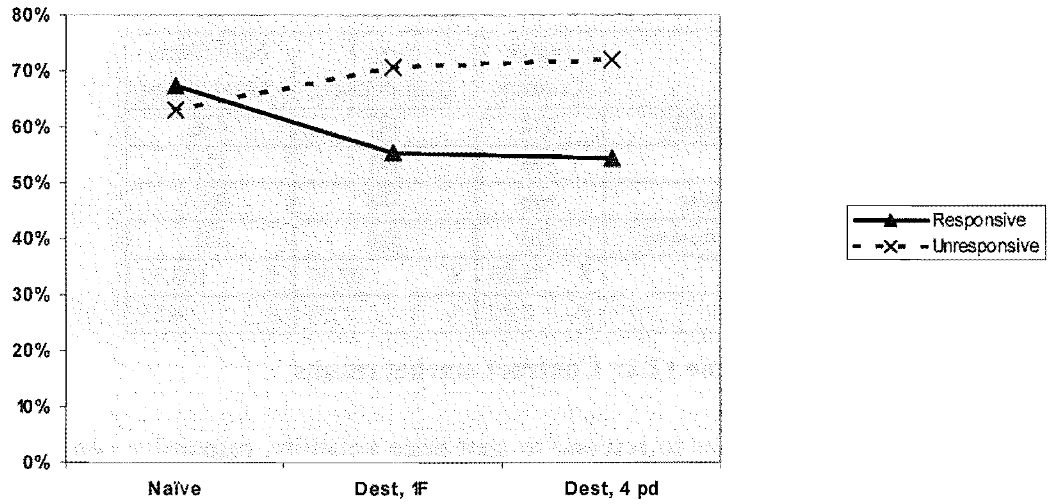


Figure 12.7 Contracts as a percentage of average load

The mean-variance maximising incentives, acting on these firms, leads them to respond to pseudo-variance in opposite ways. Responsive customers choose to bear the cost of load response as spot volatility increases, by purchasing proportionally fewer contracts, while unresponsive consumers, who have no alternative mechanism for hedging risk, maximise expected utility by purchasing more contracts, at a greater cost. Clearly the latter consumers are "held to ransom" by the generators to a much greater extent than their responsive counterparts.

Profit and Risk

Figure 12.8 describes the average profit, per hour, for each generator. Firstly, it is clear that, while Firm 1 has a steeper supply curve, its average marginal, and thus total, costs are significantly less than for Firm 2 at the equilibrium market quantities. This is also true of all individual states except the lowest cost state. In this state, output for each firm is high, and hence Firm 2 can produce at a lower marginal cost.

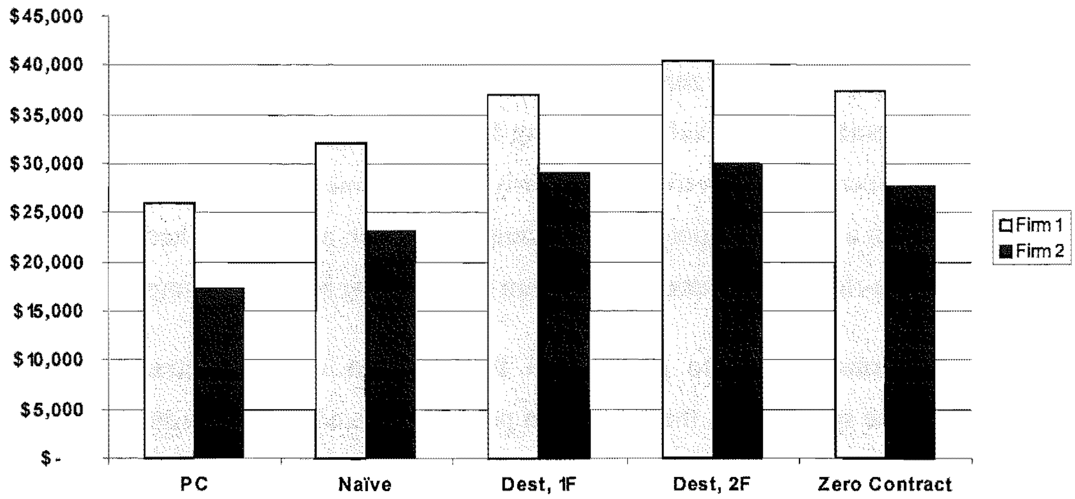


Figure 12.8 Generator profit per hour

As expected, the competitive situation delivers the lowest average profits for each firm. It is interesting to note that in the naïve case, when firms manage their spot behaviour in ignorance of the effect the spot price distribution has on contract prices, the presence of the contract market decreases profit (by around 12% for both firms) when compared with the Cournot equilibrium without contracts. While an increase in consumer risk aversion would increase contract profits in the naïve case, these results reinforce the general result that positive levels of forward contracts force generators to behave more like perfect competitors, and thus earn profits closer to competitive levels. Two effects are noted with respect to naive profits. Firstly, the higher output results in lower spot prices, and thus spot profits, than the zero-contract case. Secondly, these lower spot prices drive contract prices down, as does the stabilising effect of dominant firm behaviour (as noted in Chapter 8) via the risk premium. Hence contract prices are not significantly greater than the expected spot price, implying that using the contract market will not increase total profit significantly. In order for naïve and zero-contract profit to be equal in the scenario examined here, the consumers' degree of relative risk aversion would have to be significantly greater than its current level of between 1 and 2, in order to generate high enough contract prices for the presence of a contract market to increase profits for a naïve Cournot firm. Hence, as long as the naïve separation between contract and spot market management exists, generators may not, in fact, find it profitable to use the contract market, even when risk premiums are offered by consumers.

In the scenario where only Firm 1 destabilises, price spreading is only marginally profitable compared with the zero-contract result, and significantly less than it could earn when it is supported by its rival in spreading the price. Furthermore, when Firm 2 does not attempt to increase price variance, it clearly benefits from the actions of Firm 1. Even though it acts in a short-run spot profit maximising manner, as though it had no contracts, it is able to free-ride on the extra risk premium generated by Firm 1's destabilisation.

When both firms attempt to profit from the existence of the contract market by destabilising the spot market, the increase in profit, compared with no contract market, is not dramatic. Both firms experience additional mean profits of around 5%. However, there is a significant increase in profits in some individual states. While spot profits are sacrificed through price-spreading in high-cost states, the extra contract profits are reaped in the low-cost states, with increases of up to 10% over the zero-contract level. An even more dramatic comparison is with competitive profits. While a contract market might be introduced to regulate a dominant firm's behaviour, and decrease their profits, firms experience increases in profits of up to 80% from using the contract market to their advantage, in some states. On average, profits are 30% higher than the competitive level for Firm 1, and 40% higher for Firm 2.

While we have assumed that generators are risk-neutral, the profit variance they experience is worthy of note. Their attempt to increase risk for consumers comes at a cost in terms of their own "risk", even if they are not averse to it. Destabilisation results in a near doubling in profit variance for Firm 1, and a 35% increase for Firm 2, exceeding the levels that each firm would experience in perfect competition.

Figure 12.9 illustrates the effect of the various strategies employed by the generators on responsive consumers' profits, and unresponsive consumers' costs. In order to make interpreting the graphs easier, we have assumed that unresponsive consumers earn total revenue of $80,000, which is invariant under spot market outcomes (and hence would disappear from a mean-variance first order condition, thus not changing their contact demand). Hence we can plot profits for both types of consumer.

Not surprisingly, those scenarios leading to the highest average spot price (both firms destabilising) also result in the least average profit for the consumers.

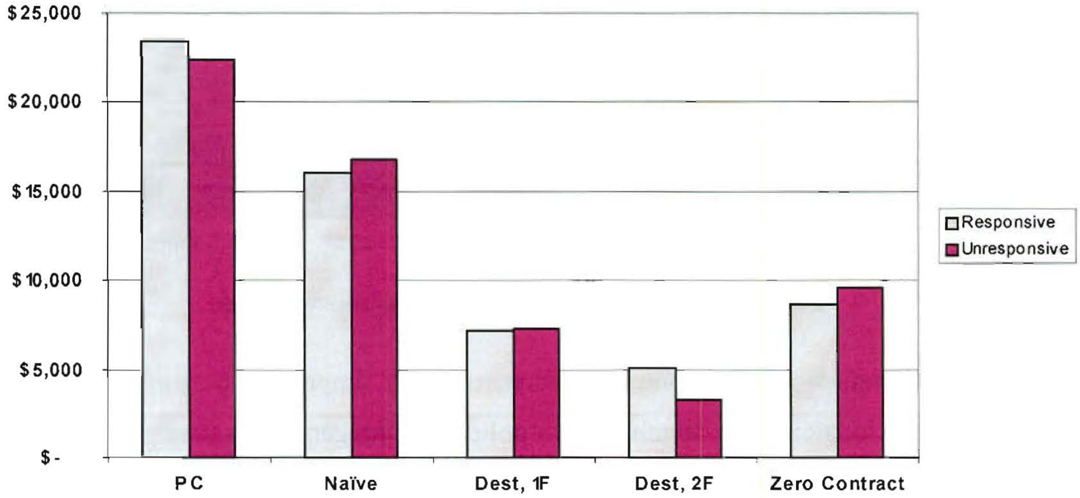


Figure 12.9 Consumer profit per hour

The effect of the generators' strategy on consumer profit volatility is clearly illustrated in Figure 12.10. It should be noted, however, that profit variance is measured given the firms' optimal contract purchases for the naïve and destabilising cases, while no contracting is accounted for in the perfect competition or zero-contract Cournot case. These figures are included to show (a) that profits are stabilised by firms utilising their market power in the spot market, ignoring contracts, and (b) that the introduction of a contract market has a positive effect on the consumers' net risk position.

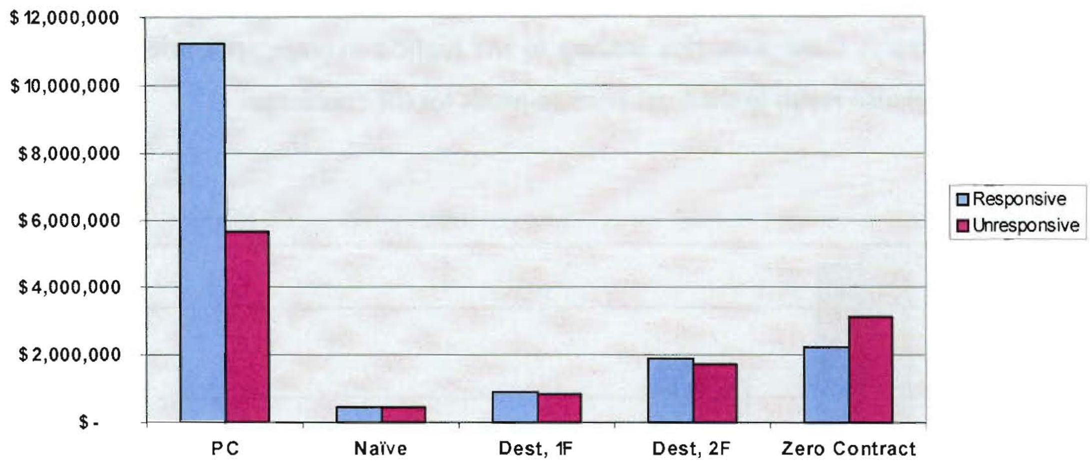


Figure 12.10 Consumer profit pseudo-variance

While the introduction of a contract market, of itself, improves the profit risk to the purchasers of electricity, the destabilisation policies of the generators leads to the greatest degree of risk when a contract market is in place.

Notwithstanding the effect of the activities of the generators on risk, it is the mean-variance objective function that dictates whether the consumers are, overall, better or worse off. It is not relevant to talk about the consumers' utility in the competitive and zero-contract case, since they do not have the opportunity to choose the level of contracts which maximises utility. Figure 12.11 illustrates the objective function value for each consumer, given their optimal choice of contracts. Not surprisingly, the combination of high and volatile prices, as both firms destabilise, results in the lowest function value for both types of consumer.

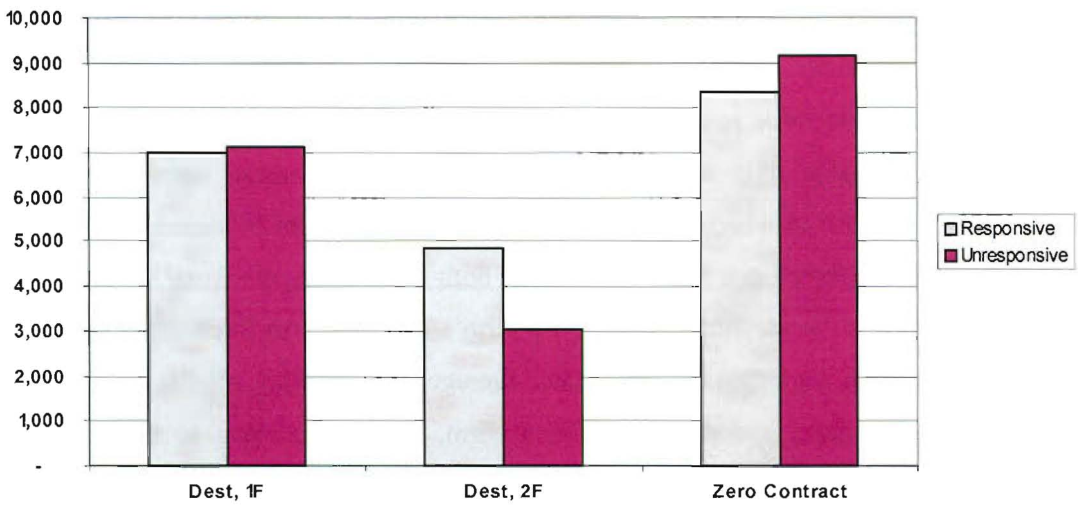


Figure 12.11 Consumer objective function

12.3.2 Independent Distributions

In the case where firms face completely independent cost-state distributions, we have chosen to model 4 states for each firm, as shown in Table 12.3.

	Firm 1		Firm 2	
State	Prob	Effect	Prob	Effect
1	0.2	0.7	0.1	0.5
2	0.3	0.9	0.4	0.75
3	0.3	1.1	0.4	1.25
4	0.2	1.3	0.1	1.5

Table 12.3: Independent Cost Distributions

These probabilities were chosen so that the underlying cost variability (measured as the variance of the average marginal cost in each state) was similar to the previous example with correlated cost states. However, comparison with the earlier results will be limited, since there are now a greater number of states, each with a lower associated probability, making it difficult to compare.

Firm 2 now faces a greater potential cost effect, although the most extreme states have a smaller probability than for its rival.

A complicating factor with independent cost distributions is the firm's choice of when to destabilise. While we assume that the firms are still aware that consumers only consider states in which the price is above the median to be risky, thus making destabilisation potentially profitable, this no longer implies *a priori* that these states will always correspond to those in which the firm's marginal cost intercept is higher than average. The spot price observed in a state is a result of both firms' marginal costs[83], which are no longer assumed to be identical. Even if the firm's marginal cost intercept is high, if its rivals cost state is significantly lower than average, the resulting market price may be lower than the mean, and the destabilising firm, acting according to the first order conditions developed in Chapter 11, will face incentives to push the price down.

Figure 12.12 shows the distribution of prices resulting from these assumptions. The states have been ordered by the number of firms destabilising. In the first four, neither firm has higher-than-average costs, and thus neither destabilises. In the next eight states, only one firm is destabilising: in the first four, it is Firm 2, and in the second four, Firm 1. In the final four states, both firms destabilise.

It is clear that in states (3,1), (3,2), (4,1) and (4,2), when only Firm 1 faces high costs, the price in the zero-contract scenario is below the median price, and thus not risky to the consumers. Furthermore, they are below the mean, and hence a firm attempting to destabilise will push these prices further down.

[83] A good indication of a high-price state would be the average marginal cost intercept in the state. However, the actual marginal cost at the profit maximising output level is a function of both the intercept and the slope of the marginal cost function. Accurate prediction of risky states prior to solving the model is not a critical issue, however, given the algorithm detailed on page 271.

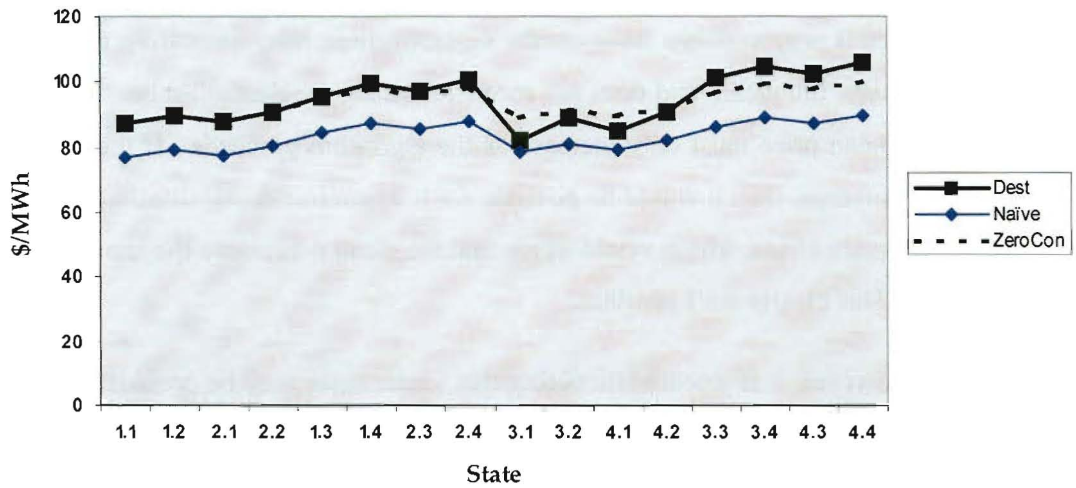


Figure 12.12 Spot price under high cost state destabilising

A potential method to achieve consistency between destabilising states and risky states is for the firm to solve the model iteratively, i.e.:

1. Solve the zero-contract model, and classify states with prices above the median price as 'risky'

2. Set both firms to destabilise in 'risky' states.

3. Observe the price distribution, and identify any states which now have prices below the median

4. Identify, for those states, which firm is driving prices down, and set that firm to not destabilise in those states.

5. Re-solve model, and return to 3.

This process repeats until consistency is found between the states that have prices above the median price, and those that would be assigned as "risky" by the consumers. The states most likely to have prices pushed down, under destabilisation, are those with short-run profit maximising prices closest to the mean, i.e., the lowest cost, risky state. Hence returning that price to its short-run value would increase the mean, and possibly increase it to a level that would induce the firm to push down the price in the next-highest cost state. This raises the question of whether the algorithm would eventually converge. However, since it progressively removes the destabilisation incentives from those states with prices below the mean, and does not reconsider them for destabilisation once it has done so, the mean price must only increase as the algorithm proceeds. If the algorithm were not to converge, then it would be possible for it to remove destabilisation incentives from all high-price states, which would imply that the mean was above the short-run level of all prices. This clearly isn't possible.

As a result, however, it is worthwhile noting that some states may be considered "risky" by consumers, but not destabilised by the firm. Note, also, that this is an algorithm that would be used by the firm, rather than the consumers. While the consumers simply assign risk to those states above the median, the generators must decide (a) which states these will be in equilibrium and (b) which of these states can be profitably spread.

This algorithm produced the price spread illustrated in Figure 12.13. When both firms experience state 3, the resulting zero-contract price was below the mean, but above the median. This implied that while the state was risky to the consumers, it was below the mean and thus driven down under the destabilisation incentives facing both firms.

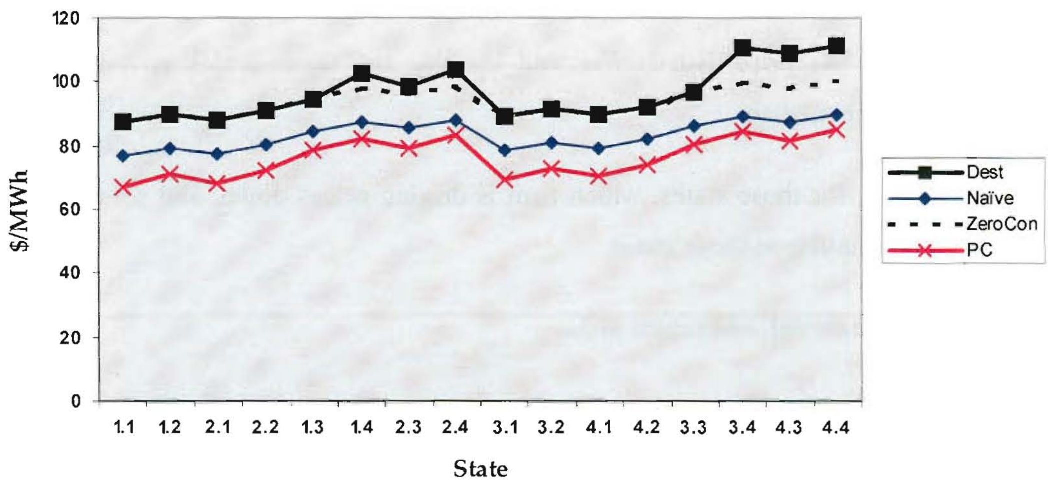


Figure 12.13 Final price spread, independent costs

While the manner in which firms find a consistent destabilisation strategy is slightly more complicated for the independent cost scenario, the resulting trends of variance, contract price and contract strategy, across the various policies are extremely similar. This should come as no surprise, since the incentives acting on the firm are identical to those in the correlated case. As a result, firms achieve similar improvements in profits by pursuing destabilisation strategies (Figure 12.14), and consumers experience a similar deterioration in their objective (Figure 12.15).

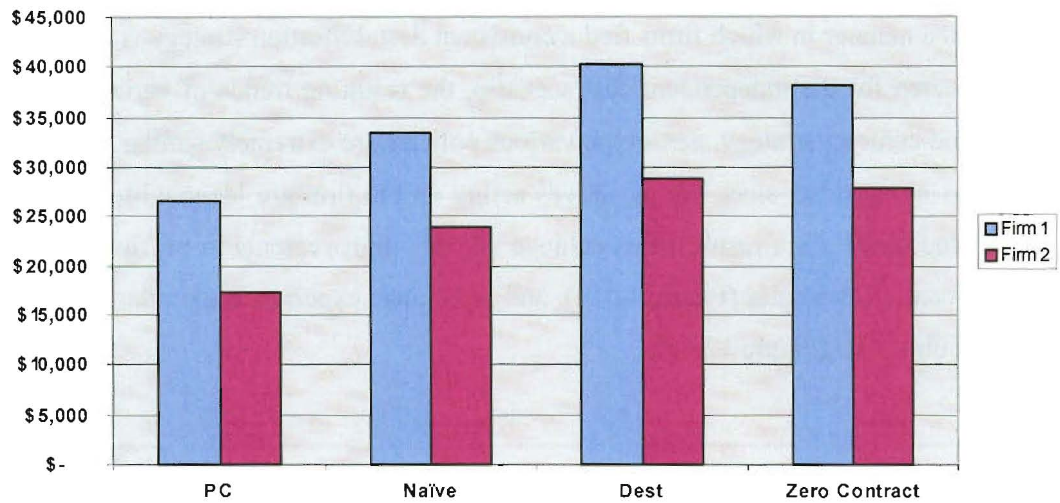


Figure 12.14 Generator profit, independent costs

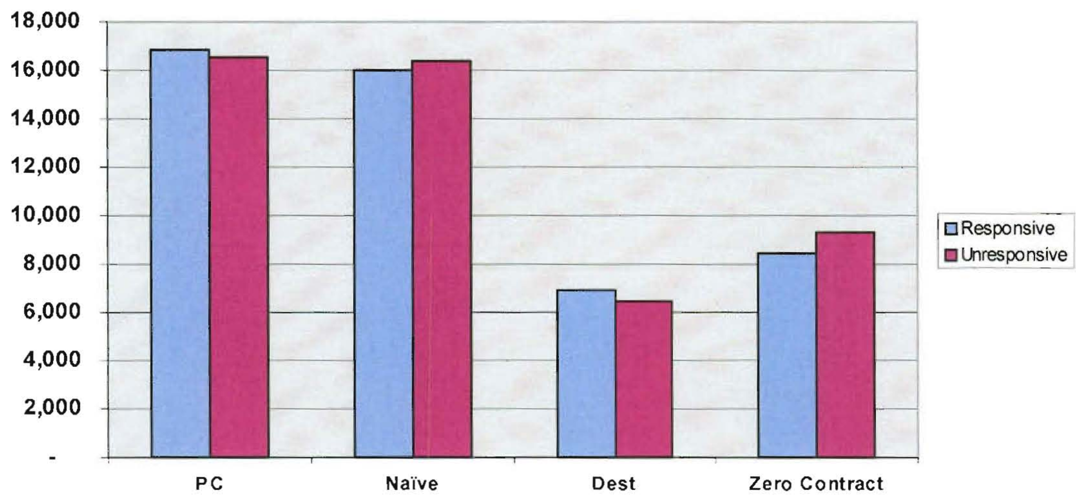


Figure 12.15 Consumer objective function, independent costs

The objective function decrease is again largely a result of both an increase in electricity costs (through the mean spot price and contract price), and an increase in the volatility of the spot price (Figure 12.16)

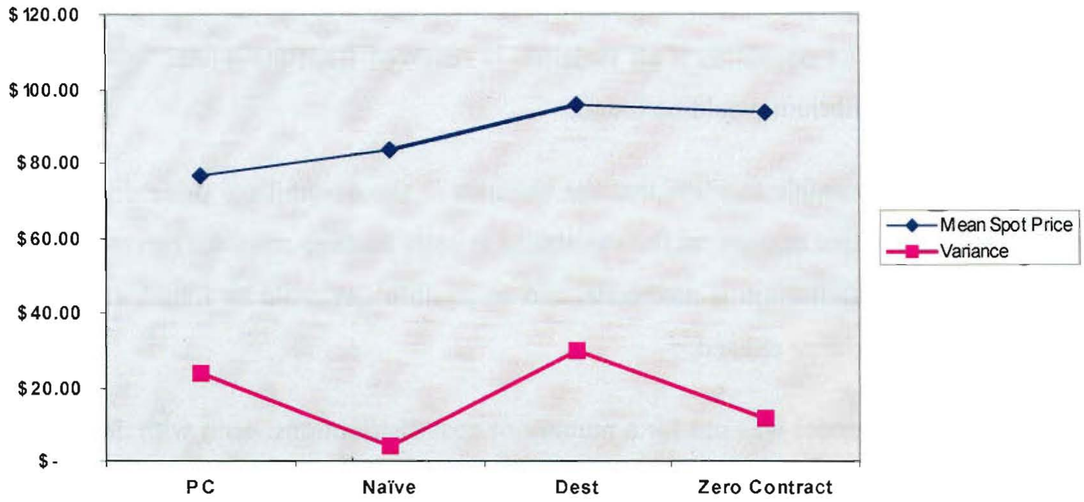


Figure 12.16 Mean and pseudo-variance of spot price, independent costs

The only significant difference to the trends noted for the correlated cost case is for Firm 2. Due to its volatile cost function, Firm 2 experiences highly variable profits. The standard deviation of its profit, over the 16 states, is as high as 30% of its average profit.

12.4 Effect of Cost Variability on Destabilising Strategy

The original motivation for a model of destabilising behaviour was that the spot market experiences variability driven by volatility in generators' cost functions. This created a "natural" level of risk in the market, which could be amplified unnaturally, by the generators varying their short-run spot behaviour, so as to create more risk for consumers and thus increase contract profits through higher risk premiums. This price spreading behaviour comes at the cost of short-run spot profits, since the firms are no longer operating at their short-run profit maximising output level.

Hence the variance in the market has, in a sense, been "seeded" by the cost variability. The question arises as to whether destabilisation is attractive, and equilibria can be found, when there is no underlying cost variation, i.e., would a firm find the "cost" of varying its spot behaviour outweighed by contract market profits.

The Cournot equilibrium model, formulated in Chapter 11, will not find this equilibrium, even if it does indeed exist. It is easy to show that the first-order conditions "collapse" to the standard Cournot equations if all variation is removed from the model, and hence a zero-variance equilibrium would be found.

However, if it is possible to show that the variance in the destabilised price distribution tends to some positive amount, as the variability in costs tends to zero, we can reasonably conclude that destabilisation is profitable, and an equilibrium could be found, even if no underlying cost variance existed.

The destabilising model was run for a number of cost distributions, each with decreasing variability. The least variability that could be obtained, before numerical difficulties were encountered, was a cost pseudo-variance of $\$0.20^2/MW^2$, compared with the cost pseudo-variance used in the results above of $\$10.00^2/MW^2$.

The spot price pseudo-variance, created by the destabilising generators, is illustrated in Figure 12.17. Clearly, it is a non-linear relationship, but does not appear to tend to zero. This would seem to indicate that with infinitesimally small variance in the underlying cost function, the firm still faces incentives to amplify the variability, in the spot price.

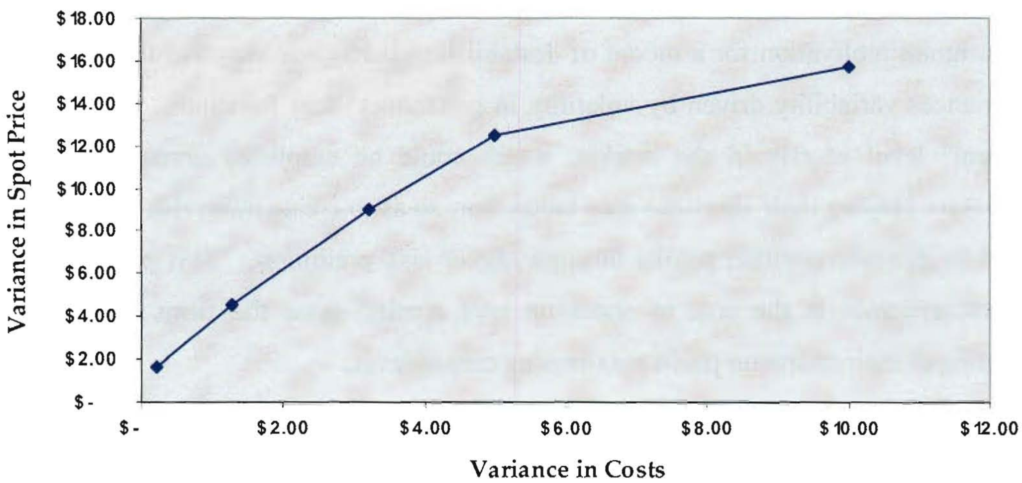


Figure 12.17 Cost and spot pseudo-variance relationship

While this does not conclusively prove that the generators could find destabilising equilibria in the absence of any "seeding" variance, we believe it motivates the formulation of a stochastic gaming model, as discussed in Chapter 11, to investigate this aspect of spot-contract dynamics.

12.5 Effect of Consumers' Risk Aversion on Destabilising Strategy

Since the extent of the consumers' risk aversion will dictate the size of the risk premium they are willing to pay for contracts, we would expect to see the degree of price spreading increase as λ increases. In fact, this was reflected in the first order conditions developed in Chapter 11, where the terms dictating the direction of deviation, from the short-run profit maximising output, were multiplied by λ. In this section, we will illustrate the response of destabilising behaviour (and profitability) to this parameter. Given the intangible nature of risk aversion, and the inability to fully represent consumers' attitudes to risk in a single parameter, it is perhaps more important to study the response of the model to changes in λ than to interpret the numerical results, for a given risk aversion value, in isolation. This is particularly important, as we are assuming an underlying quadratic utility function for consumers, which is renowned for its unusual behaviour in certain situations (e.g., decreasing utility at high levels of wealth). Despite this, the mean-variance model has great intuitive appeal, and, combined with our measure of pseudo-variance (defined in Chapter 11), we expect provides an adequate reflection of the decision making behaviour of consumers, for a given value of λ.

For simplicity, we will restrict our attention to the case where both firms destabilise, and have correlated costs. The results presented in Section 12.3 assumed a relative risk aversion corresponding to a value of λ of .0002. By varying this parameter, we have obtained results for a variety of risk aversions.

Figure 12.18 shows that the degree of price spreading achieved by the firms, as measured by the pseudo-variance, increases with the risk aversion displayed by the consumers. For very low levels of risk aversion ($\lambda = 1 \times 10^{-6}$), the pseudo-variance can be seen to approach $\$5.50^2/\text{MWh}^2$, the price variance under the zero-contract Cournot scenario.

This reflects the obvious fact that firms do not find price spreading profitable when consumers are virtually risk neutral.

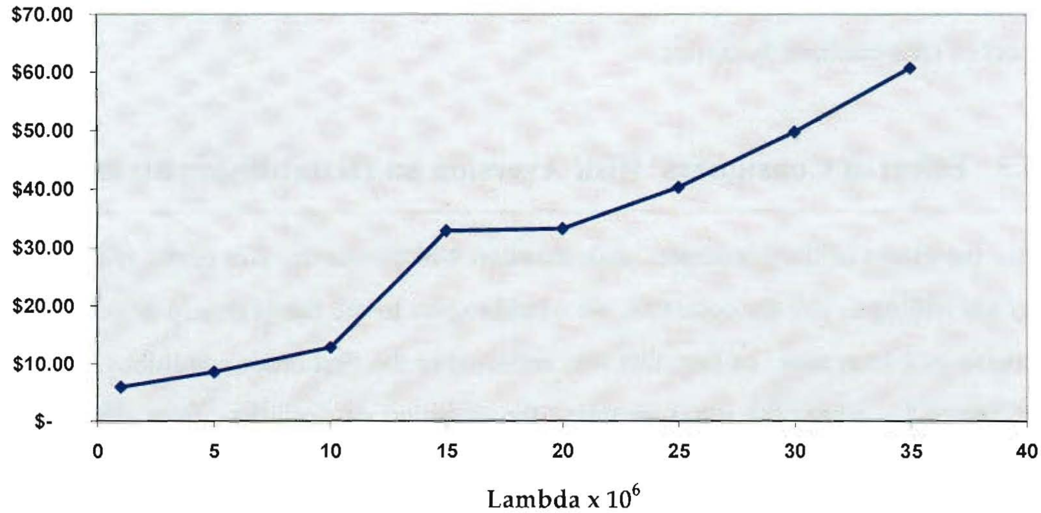


Figure 12.18 Spot price pseudo-variance under changes in risk aversion

The relationship between risk aversion, and the destabilisation achieved by the generators is by no means a smooth one. In particular, pseudo-variance at λ = .00015 seems disproportionately high, a matter we will investigate below. First, however, let us investigate the effects of the increased variance on the contract price, and optimal contract quantity sold by the generators. Figure 12.19 shows the effect of the increased destabilisation on the mean price and contract price. It should come as no surprise that the combined effect of an increasing mean, and variance, with increasing risk aversion, is for the contract price to increase approximately quadratically, and somewhat smoothly compared with the mean and variance (largely due to the high variance, at λ = .00015, being combined with a relatively low mean at that value).

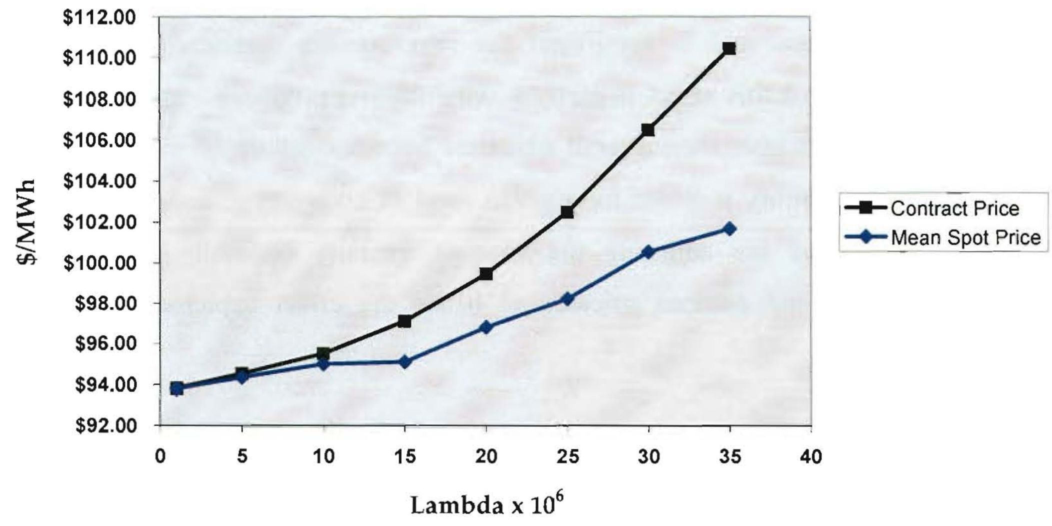


Figure 12.19 Contract price and mean spot price under changes in risk aversion

These changes in the equilibrium contract price have a much more dramatic effect on the consumers' objective function (Figure 12.21) than on generator profit (Figure 12.20). Generators' profits increase by up to 12%, from the risk neutral case, when consumers are significantly risk averse.

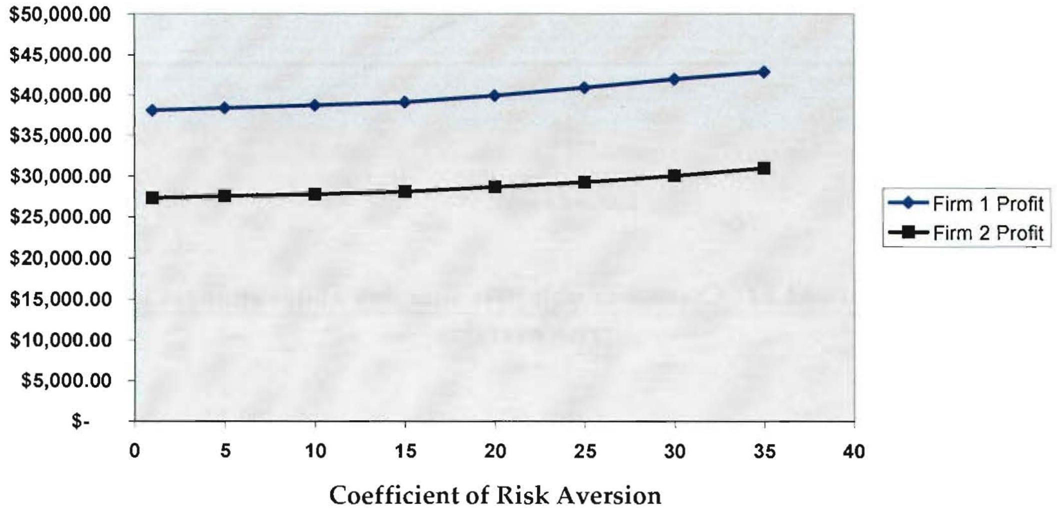


Figure 12.20 Generator profit under changes in risk aversion

However, the contract-optimised consumers' objective function is roughly halved for the responsive consumers, and is significant for unresponsive consumers also[84]. It is important to note that this effect is partly a self-fulfilling prophecy - as the consumers become more risk averse, their overall objective becomes relatively more sensitive to variance, and thus utility is lower for a given level of contracting, and given variance. However, consumers are adjusting the contract quantity optimally in response to changing variance and contract prices, and hence the effect depicted above is still significant.

As they become more risk averse, the optimal quantity of contracts purchased by the unresponsive consumers responds quite differently to that of responsive consumers. Unresponsive consumers reduce their contract purchases as they become more risk averse, while responsive consumers purchase a greater amount on contract.

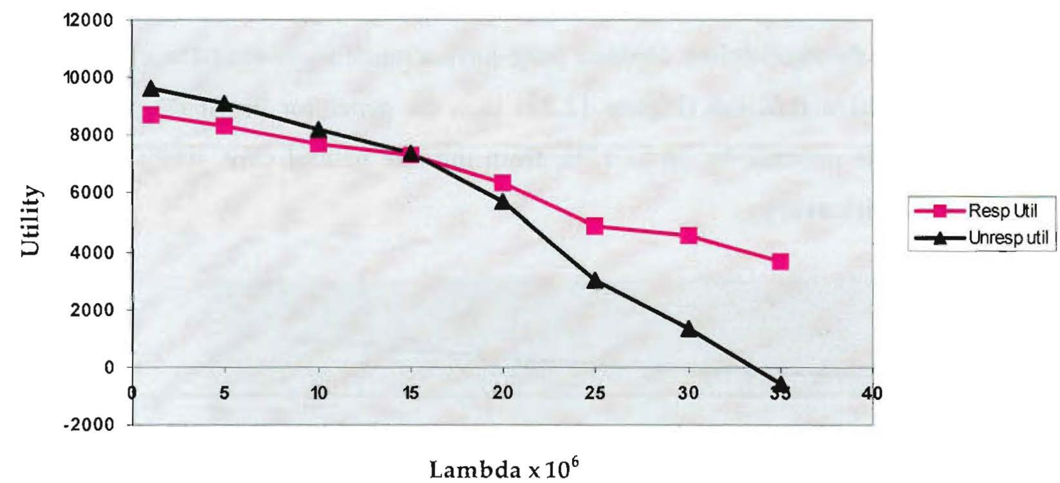


Figure 12.21 Consumer objective function under changes in risk aversion

84 The figures show that unresponsive consumers' utility becomes negative at high levels of risk aversion. However, recall that in scaling the unresponsive consumers' utility (based on electricity costs) to be comparable with that of responsive consumers, we assumed they received revenue of $80,000 in all states. Hence whether or not unresponsive utility becomes negative is sensitive to the revenue figure assumed, and it makes more sense to investigate absolute changes in utility, which, over the range of risk aversions evaluated, is approximately 4,400 utils for responsive consumers and 10,000 utils for unresponsive consumers.

There are two important effects at work in consumer contract demand, which are governed by the maximisation of mean-variance preferences: profit maximisation and variance minimisation. The generators, who are attracted to greater degrees of destabilisation as risk aversion increases, are driving up contract prices significantly, as we saw above. So while consumers may desire more certainty under higher levels of risk aversion, their mean-variance preferences provide incentives to balance risk-avoidance with profit maximisation, the latter leading them to avoid the high contract prices by reducing this 'profit-making' component of contract demand. This is true for both types of consumer, and since it is the only component of contract demand varying for unresponsive consumers (since their load is fixed regardless of the shape of the spot price distribution or risk aversion), we observe their contract demand decreasing as risk aversion increases (illustrated in Figure 12.22). Thus we can conclude that the profit-making aspect of these consumers' contract demand dominates the risk-averting aspect.

However, responsive consumers have two other varying components of contract demand. First, as the mean spot price increases (as risk aversion increases), responsive consumers reduce their total electricity load. Secondly, the "natural hedging" component of their demand, reflecting the covariance of their profits with the electricity price, will also change as prices become more variable. Given that it is difficult to intuitively explain the relationship between the variance and covariance of a distribution, we can only numerically observe changes in response to the generators' increased destabilisation. Since these consumers increase their optimal contract purchases as risk aversion increases, it is clear that the behaviour of this contract demand component dominates the effect of the profit-making term, and their reduced electricity purchases. We can conclude that the distribution resulting from the generators' destabilisation at higher levels of risk aversion are such that the natural hedge provided by their load response is less effective.

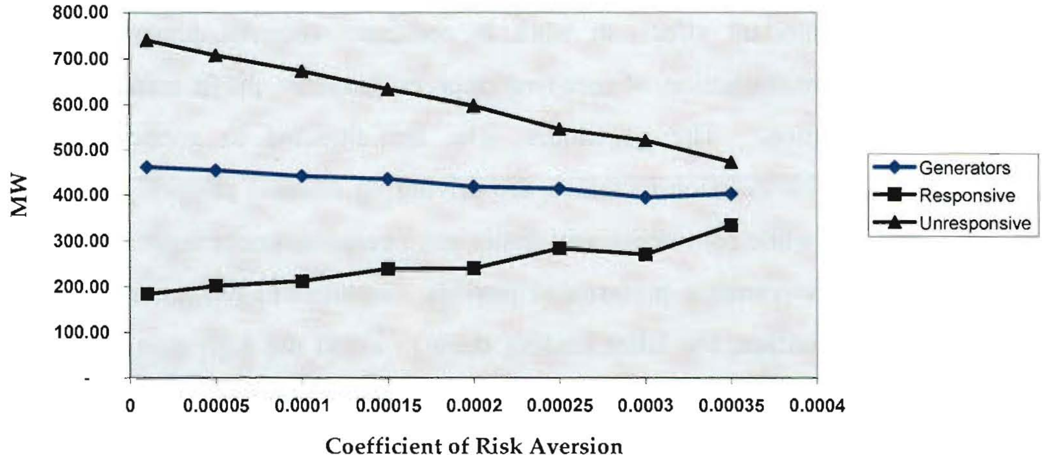


Figure 12.22 Contracts traded in equilibrium, under changes in risk aversion

We will return now to the anomaly highlighted above, namely that the pseudo-variance, when λ = .00015, seemed disproportionately large compared to its neighbours. Figure 12.23 illustrates the way in which the generators are increasing the spread of prices, as risk aversion increases, and shows that it does not always increase the price in every high cost state. For λ = .00015, and in fact λ = .0002, prices in state 5 are pushed down.

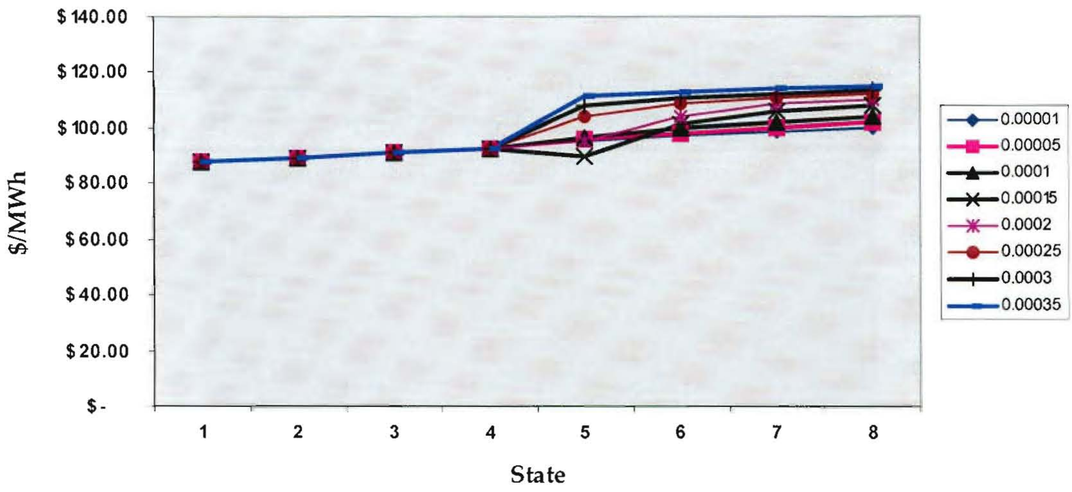


Figure 12.23 Spot price spread for various risk aversions

This is an identical effect to that noted for the independent cost scenario (Section 12.3.2). As the mean increases under price spreading, prices in states which have a cost close to the mean cost, require a greater degree of spreading to achieve a profitable effect on the pseudo-variance, particularly at low levels of risk aversion. The decrease in generation required for an "effective" spread may be too costly in terms of sacrificed spot profits. The first order conditions, in themselves, made no distinction between risky, and non-risky states, and simply led the firm to push the price away from the mean, in whichever direction was most profitable. Hence in state 5, the firms instead increase generation to force the price below the mean, even though it was a state identified by consumers as being "risky". This introduces a possible inconsistency in the model, in that states that are, *a priori*, high-priced and thus risky for consumers, may become low-priced states, and not of concern to the consumer any more. Section 12.3.2 proposed an algorithm to ensure that there is consistency between risky states prior to, and after, destabilisation. Figure 12.24 illustrates the price spread resulting from a single iteration of this algorithm, compared to the original spread from Figure 12.23.

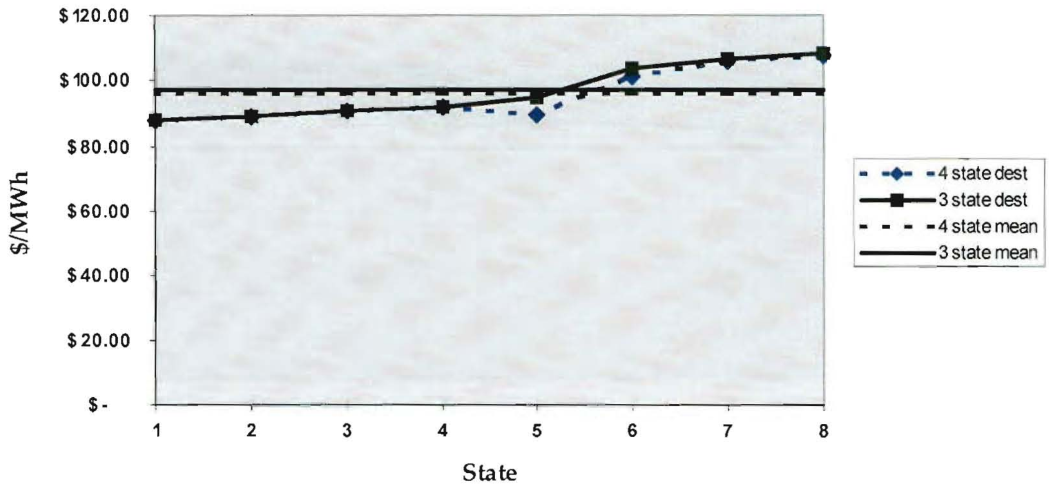


Figure 12.24 3 vs 4 destabilising states

It is clear from Figure 12.24 that the zero-contract solution lies below the mean in both cases, hence the response of the firms to push it down when destabilising. It is also clear from the figure that a lower degree of pseudo-variance is now observed, since the price in state 5 is now very close to the mean. The pseudo-variance is now $\$27.57^2/\text{MWh}^2$, compared with $\$32.92^2/\text{MWh}^2$ measured in the model when the price was pushed down; however, in that case, consumers would not have considered the state risky, and an overall pseudo-variance of $\$26.50^2/\text{MWh}^2$ would have been measured by the consumers. Hence the generators have gained from re-solving the model.

These results provide a sensitivity analysis with respect to a somewhat arbitrary parameter, the coefficient of relative risk aversion. The behaviour of the firm, or at least the absolute variance achieved from destabilisation, does appear relatively sensitive to the consumer risk aversion. However, profit does not seem overly sensitive, reflecting the tradeoff between the returns to destabilisation, through contract profits, and the implied cost of moving away from the short-run profit-maximising output level.

12.6 Conclusions

The aim of this chapter was to show that, given some reasonable parameter values and cost state distributions, firms could find equilibria in both the spot and contract market that maximised the combination of spot and contract profits. Stable equilibria were found both for the case where firms did not anticipate the effect spot outcomes had on contract prices, and when they did.

The two LRE models, developed in the previous two chapters, were formulated in GAMS, and solutions found. In both models, consumers only viewed those states with prices above the median price as "risky", and thus only these states were worth destabilising to increase their assessment of risk and thus contract prices. Destabilising equilibria were sought and compared with scenarios in which the firms acted as perfect competitors, and Cournot competitors, both in the absence of a contract market.

These results were intended to be generally indicative of the incentives for dominant firms, with access to risk-averse consumers and a contract market, to purposely amplify the natural variability in the spot price that results from their own variations in costs. Usually, contract markets are introduced to induce firms with a high degree of market power into producing closer to the competitive equilibrium. They also provide a mechanism for risk-averse consumers to purchase fixed-price contracts, in order to hedge the risk they face in their own markets.

However, the above results show that if generators are able to extract risk premiums from these consumers, and the size of the premium is driven by spot price variance, generators can use destabilisation strategies to deliberately shift the contract demand curve upwards, to their advantage. Thus, even though consumers continue to contract optimally by maximising a mean-variance function of wealth, the actions of the generators make a worse utility position unavoidable for the consumers. When dominant firms jointly optimise the contract and spot market strategy, taking account of the relationships between them, they increase the variance of the spot price to levels which are greater than in the competitive case, and this contrasts strongly with the conventional Cournot result, in which risk is sharply decreased. Furthermore, profits to the generators are significantly greater than the competitive case, although, interestingly, not significantly increased from the zero-contract Cournot level.

Results were obtained both in the case where generator cost functions were perfectly correlated, and when they were completely independent. In both cases, firms obtained similar improvements in profits from destabilisation. However, in the independent cost case, firms found that experiencing a high-cost state did not imply *a priori* that destabilisation was profitable in that state. In some cases, where only one firm faced a high cost state, and the other a low-cost state, even the zero-contract Cournot prices would have not been considered high, and therefore risky, by consumers. Thus the incentives acting on the firm to spread prices would have pushed prices down, since the expression itself does not include a variable determining whether the state is "risky" or not. If consumers do not consider low prices to be risky, pushing prices down is of no value to the firm in the contract market, since it would not actually increase the risk premium offered by consumers. This scenario could also occur if the price-spreading behaviour of the firm raised the mean price sufficiently to change the incentives in some states to be price-depressing, rather than price inflating. An algorithm was introduced that pursued consistency between the states viewed as risky by the consumers, and those in which destabilisation was pursued by the generators.

We also showed that even in the event of almost zero variability in costs, the firms still found equilibria that supported an amplification of the spot price variance. This would indicate that destabilising strategies may be sought even by firms that have complete certainty about their costs. While this is not possible in the model proposed here, in which destabilisation is "seeded" by cost variation, we believe this motivates the investigation of a stochastic model which employs randomised strategies to find equilibria under the zero cost variability scenario.

Finally, we showed that while the degree of consumers' risk aversion had a significant effect on the degree of price spreading pursued in equilibrium, and the resulting contract prices, it did not have a dramatic effect on total profits, due to the optimal contract demand response from consumers. As risk aversion increased, destabilisation appeared more attractive, since risk premiums were greater. However, since consumers' utility functions led them to maximise a combination of profit and variance of profit, the higher mark-up of contract prices over expected spot prices meant consumers found the gamble of the spot market more attractive.

13

CONCLUSIONS

13.1 Introduction

This thesis has addressed aspects of the risk management problem faced by the manager of a large hydroelectric generation firm in a deregulated electricity market. This required a synthesis of three important areas of modelling in electricity markets: the management of storage facilities, the use of market power by dominant firms, and the use of financial contracts to hedge risk. It was proposed that these mechanisms should not be considered in isolation, and that a model intending to optimise the behaviour of a supply firm in this situation must include as many of the interactions between these as possible. A "risk management triangle", illustrated again in Figure 13.1, was used to represent the framework we applied to the development of the ideas presented in the thesis.

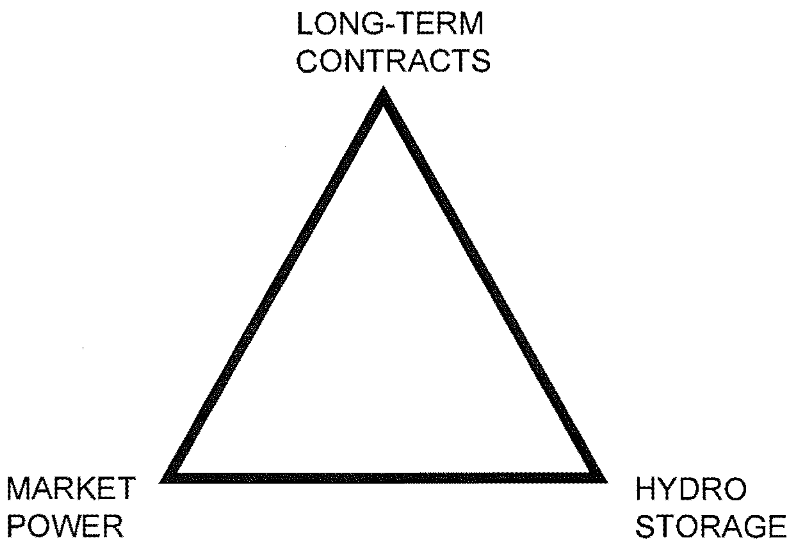


Figure 13.1 The Risk Management Triangle

This chapter summarises the main results and ideas presented in this thesis, and outlines possible areas for future study. Section 13.2 outlines the insights that were derived about the effect of profit-maximising reservoir management policies and contracts on profit risk, for a dominant hydro generator. Section 13.3 outlines the insights and results from a model, formulated in this thesis, of the joint contract and spot problem faced by a dominant generation firm with uncertain costs.

In developing this model, a number of simplifying assumptions were made. Section 13.4 discusses the impact these assumptions might have on the results, and the implications for the behaviour that might be seen in reality. Suggestions for further research are presented.

13.2 Contracts and Risk

While uncertainty is prevalent in all electricity markets, hydro generation introduces an extra dimension of volatility. In order to investigate the behaviour of the firm under uncertainty in hydrological inflows, and its effect on firms' profit variability, we used an existing model of a Cournot duopoly. The model consisted of a hydro and a thermal firm, and both were assumed to be risk neutral. The quantity of contracts sold by the firms was fixed, and the contract price was equal to the mean spot price. We found that even for a relatively high degree of inflow variability, the dominant hydro firm was able to utilise both its reservoir, and its market power, to the extent that the variance of profit was low, under a range of simulated inflow sequences. If profit variance is to be taken as our measure of risk, these results suggest that including risk aversion (if appropriate) in a model of a dominant hydro firm with sufficient hydro storage would not significantly change the firm's behaviour. As long as the firm was managing its storage position in a profit maximising manner, profit variability would remain at levels that would be acceptable for most managers, whether or not they were risk averse.

The results also showed that increasing the level of contracts sold by the hydro firm significantly decreased mean profit, and moderately increased the level of risk. Contrary to the conventional wisdom that forward contracts provide more certainty, the latter observation supports recent analyses of electricity markets which suggest that selling more contracts can increase risk for firms that are unsure of their ability to generate electricity. This is not unique to hydro systems, since factors such as plant and transmission failure, and the uncertain availability of generation fuels exist in the vast majority of electricity markets. In our simulation, the greatest levels of profit variance were experienced by the hydro firm at high levels of contracting. However, even at these levels of contracting, the standard deviation of mean profit did not exceed 10% of the mean level.

This analysis showed that, at least for the particular market setting[85], the three corners of the risk management triangle combine to yield a low level of profit risk for the generator. If a mean-variance model reflected the preferences of the hydro manager, the optimal level of contracts would be low, if not zero, according to the numerical results from this model. Given that a firm, acting in a profit maximising manner, could adequately manage risk through the use of its market power and reservoir, the only reasonable incentive for it to sell long term contracts would be to increase profits. As stated above, as the firm increased its level of contracts, its profit maximising output increased, and thus spot prices decreased, relative to the zero-contracting Cournot output. In this model,

[85] Although, as noted in that chapter, similar results were observed for a moderate range of market situations, including constant elasticity competition and a variety of storage capacities.

as is the case in many futures markets analyses, the contract price was set equal to the mean spot price. Hence contracts committed the generator to a lower profit level than that which is observed when a contract market is not present. In order to explain why generators might sell positive levels of contracts, we suggest that contract prices may be greater than the expected spot price, i.e., contract prices would include a premium determined by the level of risk observed by, and the risk aversion of, electricity consumers. This allows generators to sell output in the contract market at a premium to what they would receive in the spot market.

To reflect this, a model of contract demand was developed, and mathematical expressions obtained for optimal contracting, by electricity consumers who had mean-variance risk preferences. Two important classes of consumers were investigated: large industrial consumers, who respond to the electricity price by altering their level of activity and thus electricity load, and electricity retail firms, who have a highly inelastic load representative of most small households. The analysis showed that while risk averse unresponsive consumers will certainly offer premiums for contracts that hedge spot price variance, responsive consumers may not (to the same degree), since they may choose to alter their load in response to the spot price. Since the response has a stabilising effect on overall profit variability, these consumers will be less inclined to pay a premium to purchase fixed-price contracts.

13.3 Joint Spot and Contract Equilibria

Assuming that we could ignore risk aversion on the part of the generators, and that storage is managed optimally by those firms which have access to it, the problem facing the generators reduces to deciding optimal levels of contracts and spot output. In order to reflect the natural variability that would drive contract demand, generators were assumed to face uncertain costs (interpreted, for the hydro firm, as variable marginal water values arising from the optimal reservoir management policy).

We proposed a concept of long-run equilibrium, which, while hypothetical, had convenient properties that would make analytical joint spot- and contract-profit maximising conditions for the problem faced by the firms relatively tractable. The full set of profit maximising conditions reflected the following decision structure:

- Generators faced a discrete distribution of states, corresponding to different levels of marginal costs, and determined profit maximising output levels, and thus Cournot-Nash equilibrium spot prices.

- This gave rise to a distribution of spot prices, which, when observed by consumers, determined their optimal contract demand curve.

- Generators would compete on this demand curve, *á la* Cournot again, in the contract market.

- Since the contract market occurred prior to the spot market, and to the resolution of cost state uncertainty, a single contract level was sought that maximised expected profit over all cost states.

Given that we were aiming to find **joint** contract and spot optima, the way in which generators managed each market, and the information passed between them, was important. Two models were proposed. The first model suggested a separation between contract and spot decisions: a spot manager sets the profit maximising output assuming the contract level and price is fixed, with the contract level being optimally set as a function of the resulting variance and covariance of the distribution of prices. This gave rise to extended versions of equations developed previously by a range of authors. An increase in the level of contracts leads to higher output, and lower prices, since the spot manager does not account for the effect the market outcomes have on contract profits, in equilibrium.

A second model was proposed, in which the spot manager was aware that his/her choice of output level would have a direct impact on the distribution of prices, and thus on contract demand. The greater the variance, or the lower the covariance, created in the spot market, the higher the contract prices offered in the contract market. First-order conditions were developed that intuitively reflected this notion, and three important effects were noted:

- Generation was no longer strictly increasing in the contract level. In fact, firms "anchored" their generation decisions in each state to the zero-contract output, acknowledging the fact that while increasing generation would achieve greater spot profits, it would decrease contract profits through the effect it had on the mean spot price.

- Since customers with unresponsive loads were averse to the variance of the spot price, and would offer increased contract premiums for greater levels of spot variance, an incentive existed for generation firms to spread prices, or deliberately "destabilise" the market: in states that naturally led to prices above the mean (i.e., high cost states), prices were forced higher, while in those states where costs were low, prices were driven lower.

- Simultaneously, customers who had the flexibility to respond to spot variation by altering their electricity purchases placed downward pressure on contract prices. This effect was largely represented in the contract demand curve by the covariance of the spot price with its square. Since the covariance decreased risk premiums on forward contracts, generation firms faced an incentive to stabilise prices, relative to those resulting from a zero-contract Cournot policy.

Which of the latter two effects dominated, in aggregate, determined whether firms would ultimately find it profitable to stabilise or destabilise the spot market.

Notwithstanding some difficulties in ensuring profit concavity in the second model, results were obtained for each model, and compared with traditional Cournot and perfect competition results, in the absence of a contract market.

Results clearly demonstrate that the way in which firms jointly manage contract and spot market behaviour has a significant impact on profit. If firms act in ignorance of the effect spot behaviour has on the contract market, then a positive level of contracts leads to higher output, on average, and lower spot prices. Despite the risk premiums available through a variable spot price, overall contract and average spot prices will be relatively low, leading to profits that are only moderately greater than competitive ones.

However, once a firm considers the effect spot behaviour has on contract prices, equilibria can be found that provide greater profits than in the zero-contract Cournot case. Furthermore, the incentives to destabilise the spot market, i.e., to spread the spot price distribution wider than in the zero-contract Cournot case, dominated the stabilising incentives in all cases investigated. Hence the short-run spot profits sacrificed in the process of destabilisation were made up for in higher contract revenue, both through an increased expectation of the spot price, and a greater variance of the spot price.

This illustrates an aspect of risk management that has been largely ignored in the literature: namely, that **generators may intend to *create* risk for other market participants, instead of, or in addition to, reducing it for themselves.**

13.4 Interpretation of Results and Further Research

Our intention, to find strategies for dominant firms which jointly optimise contracts, storage and market power, led us to develop a joint spot/contract equilibrium model. The model (and thus the results above) relies heavily on Cournot behaviour by the firms, which represents an extreme assumption with respect to the firms' ability to exercise market power, and should be viewed as giving an outer limit result. In particular, it may overstate the ability of firms to drive prices up in high-cost states, in order to spread prices. If competition was modelled more realistically (for example, using supply functions), firms may not find sacrificing short-run spot profits as attractive, especially if their efforts are not going to impact the price distribution to the same degree. In the limit, i.e., perfect competition, individual firms would not be able to affect the consumers' perceptions of price variance at all, and hence the short-run profit maximising solution would be the optimal strategy.

Equally, risk aversion could be introduced for generation firms. While our numerical analysis in Chapter 8 showed that minimal levels of risk existed for firms acting in short-run optimality, the results in Chapter 12 showed that the destabilisation efforts of the generators increased their own profit variance, as well as that of the consumers. It may turn out that, in order to profitably create risk, their own levels of risk become unacceptable.

We believe that an important addition to the model would be to include entry, especially given that we are modelling a market in which there are few firms, supporting prices well above short-run marginal cost. Entry would potentially limit the market power of the firms, in both the naïve and destabilising models. In particular, destabilisation might increase the prospects of entry, as price spreading improves the potential spot profitability of the entrant, and higher contract prices give the entrant an opportunity to secure a greater long-term, guaranteed income stream. However, given the number of factors influencing an entry decision, and the manner in which incumbent firms may respond to potential entrants, its inclusion in the model presented here was, analytically speaking, too complicated. We suggest this is an important direction for future research.

Despite these limitations, we have demonstrated an important result, that while contract markets are often introduced in the hope of inducing more competitive behaviour in dominant firms, such firms may use the contract market to their advantage, and to the detriment of consumers. The insights developed may, in fact, have wider applicability than the situation modelled here. For example, we assumed that there existed a "natural" level of variability in the market, driven by generators' cost uncertainty, and incentives existed for this variability to be amplified. In fact, this underlying variability could be caused by almost anything, and, indeed, need not actually exist. As long as generators vary their spot output, consumers will see the market as volatile and risky, and desire to enter into long term contracts to hedge their risks. Furthermore, given that the state of long-run equilibrium proposed in this model is not likely to occur in reality, consumers may never have "complete" knowledge of the range of outcomes that can occur. Hence generators may have the added advantage of "surprising" the consumers.

In summary, we believe the analysis contained within this thesis should motivate further work in this area. Many electricity markets around the world are still characterised by a few dominant firms, and notable events in recent years have highlighted the volatility of the market for consumers and generators alike. The implications for risk management strategies are wider than the traditional wisdom would suggest, as evidenced by the profitability of "risk creation" reported above.

BIBLIOGRAPHY

Aghion, P. and Bolton, P. (1987) "Contracts as a Barrier to Entry", *American Economic Review*, **77**(3), 388-400.

Allaz, B. (1991) In *Commodity, Futures and Financial Markets* (Ed, Phlips, L.), Kluwer Academic Publishers,, pp. 249-71.

Allaz, B. (1992) "Oligopoly, Uncertainty and Strategic Forward Transactions", *International Journal of Industrial Organisation*, **10**, 297-308.

Allaz, B. and Vila, J.-L. (1991) "Cournot Competition, Forward Markets and Efficiency", *Journal of Economic Theory*, **59**, 1-16.

Amir, R. and Jin, J. (2001) "Cournot and Bertrand Equilibria Compared: Substitutability, Complementarity and Concavity", *International Journal of Industrial Organisation*, **19**, 303-17.

Anderson, E. J. and Philpott, A. B. (2002a), To appear in *Decision Making Under Uncertainty: Energy and Environmental Models* (Eds, Auzerais, F., Burrage, R., Greengard, C. and Ruszczynski, A.), Springer-Verlag.

Anderson, E. J. and Philpott, A. B. (2002b) "Optimal Offer Construction in Electricity Markets", *Mathematics of Operations Research*, **27**(1), 82-100.

Andrews, C. J. and Govil, S. (1995) "Becoming Proactive about Environmental Risks", *Energy Policy*, **23**(10), 885-92.

Arrow, K. J. (1971) "The Theory of Risk Aversion", In *Essays in the Theory of Risk Bearing*, North-Holland, Amsterdam, pp. 278.

Baldick, R., Grant, R. and Kahn, E. (2000) "Linear Supply Function Equilibrium: Generalizations, Application and Limitations",Working Paper PWP-078, University of California Energy Institute, Berkeley.

Baldick, R. and Hogan, W. (2001) "Capacity Constrained Supply Function Equilibrium Models of Electricity Markets: Stability, Non-decreasing Constraints and Function Space Iterations",Working Paper PWP-089, University of California Energy Institute, Berkeley.

Bell, D. E. (1995) "Risk, Return and Utility", *Management Science*, **41**(1), 23-30.

Bolle, F. (1992) "Supply Function Equilibria and the Danger of Tacit Collusion", *Energy Economics*, **14**(2), 94-102.

Bolle, F. (2001) "Competition with Supply and Demand Functions", *Energy Economics*, **23**(3), 253-78.

Borenstein, S. and Bushnell, J. (1998) "An Empirical Analysis of the Potential for Market Power in California's Electricity Industry",Working Paper PWP 044r, University of California Energy Institute, Berkeley.

Borenstein, S. and Bushnell, J. (2000) "Electricity Restructuring: Deregulation or Reregulation?", *Regulation*, **23**(2), 46-52.

Borenstein, S., Bushnell, J. and Wolak, F. (2000) "Diagnosing Market Power in California's Deregulated Wholesale Electricity Market",Working Paper PWP-064, University of California Energy Institute.

Brennan, D. and Melanie, J. (1998) "Market Power in the Australian Power Market", *Energy Economics*, **20**(2), 121-33.

Brianza, T., Phlips, L. and Richard, J.-F. (1987) "Futures Markets, Inventories and Monopoly",Working Paper Discussion Paper 8725, CORE.

Brander, J. and Lewis, T., (1986) "Oligopoly and Financial Structure: the Limited Liability Effect", *American Economic Review*, **76**, 956 - 971

Brooke, A., Kendrick, D. and Meeraus, A. (1992) *GAMS: A User's Guide,* Scientific Press, San Francisco, CA.

Bunn, D. W., Larsen, E. R. and Dyner, I. (1997) "Modelling Latent Market Power across Gas and Electricity Markets", *System Dynamics Review*, **13**(4), 271-88.

Burchett, J. F. and Tummala, V. M. R. (1999) "Applying a Risk Management Process (RMP) to Manage Cost Risk for an EHV Transmission Line Project", *International Journal of Project Management*, **17**(4), 223-35.

Bushnell, J. (2000) "Water and Power: Hydroelectric Resources in the Era of Deregulation in the Western US", Working Paper PWP-056r, University of California Energy Institute.

Commerce (1999) *Energy Data File July 1999,* Ministry of Commerce, Wellington, pp. 39-40.

Cozzolino, J. M. (1979) "A New Method for Risk Analysis", *Sloan Management Review*, Spring 1979, 53-66

Dahl, C. (1996) "U. S. Energy Product Supply Elasticities: A Survey and Application to the U.S. Oil Market.", *Resources and Energy Economics*, **18**, 243-63.

Dalziel, P. C. (1987) "Optimal Water Storage and Pricing: The Effect of Monopoly", *New Zealand Economic Papers*, **21**, 3-16.

Drud, A. (1992) "A Large Scale GRG Code", *ORSA Journal on Computing*, **6**, 207-16.

Dyer, J. S. and Sarin, R. K. (1982) "Relative Risk Aversion", *Management Science*, **28**(8), 875-86.

Elton, E. J. and Gruber, M. J. (1995) *Modern Portfolio Theory and Investment Analysis,* John Wiley and Sons, New York.

Fehr, N.-H. M. v. d. and Harbord, D. (1993) "Spot Market Competition in the UK Electricity Industry", *Economic Journal*, **103**, 531-46.

Fischhoff, B., Watson, S. and Hope, C. (1984) "Defining Risk", *Policy Sciences*, **17**, 123-39.

Fishburn, P. (1984) "Risk as Probable Loss", *Management Science*, **30**(4), 396-406.

Fishelson, G. (1989) "Imperfect Competition Under Uncertainty", *Journal of Economics and Business*, **41**, 253-63.

Fleten, S.-E. and Wallace, S. (1998) "Power Scheduling with Forward Contracts", *Nordic Mathematical Programming Symposium*, Molde, Norway, Francis, J. C. and Archer, S. H. (1979) *Portfolio Analysis,* Prentice Hall, New Jersey.

Gans, J., Price, D. and Woods, K. (1998) "Contracts and Electricity Pool Prices", *Australian Journal of Management*, **23**(1), 83-96.

Gedra, T. W. (1992) "Calls and Callable Forwards for Electric Power Markets", *The Impact of a Less-Regulated Utility Environment in Power System Control and Security*,,

Gersten, A. (1999) "Hedging Your Megawatts", *Journal of Accountancy*, **188**(5), 47-54.

Green, R. (1993) "Long-term Contracts and the Electricity Spot Market", *RES Conference*, York,

Green, R. (1996) "Increasing Competition in the British Electricity Spot Market", *Journal of Industrial Economics*, **44**(2), 205-16.

Green, R. (1999) "The Electricity Contract Market of England and Wales", *The Journal of Industrial Economics*, **47**(1), 107-24.

Green, R. and Newbery, D. (1992) "Competition in the British Electricity Spot Market", *Journal of Political Economy*, **100**(5), 929-53.

Guan, X., Ho, Y.-C. and Pepyne, D. (2001) "Gaming and Price Spikes in Electric Power Markets", *IEEE Transactions on Power Systems*, **16**(3), 402-08.

Guthrie, G. (1998) *Financial Economics: Notes to Accompany Lectures for ECON 311,* University of Canterbury, Christchurch, New Zealand, pp. 234.

Helm, D. and Powell, A. (1992) "Pool Prices, Contracts and Regulation in the British Electricity Supply Industry", *Fiscal Studies*, **13**(1), 89-105.

Hirschleifer, D. (1989) "Futures Trading, Storage and the Division of Risk: A Multiperiod Analysis", *The Economic Journal*, **99**, 700-19.

Hobbs, B. F. (1986) "Network Models of Spatial Oligopoly With An Application to Deregulation of Electricity Generation", *Operations Research*, **34**(3), 395-409.

Hobbs, B. F. (1999) "LCP Models of Nash-Cournot Competition in Bilateral and POOLCO-Based Power Markets", *IEEE Winter Power Meeting*, New York, Hobbs, B. F. and Schuler, R. E. (1985) "An Assessment of the Deregulation of Electric Power Generation Using Network Models of Imperfect Spatial Markets", *Papers of the Regional Science Association*, **57,** 75-89.

Hoff, T. (1997) "Using Distributed Resources to Manage Risks Caused by Demand Uncertainty", *The Energy Journal*, **Special Issue: Distributed Resources,** 63-84.

Hogan, W. W. (1997) "A Market Power Model with Strategic Interaction in Electricity Networks", *Energy Journal*, **18**(4)**,** 107-41.

Hughes, J. and Kao, J. (1997) "Strategic Forward Contracting and Observability", *International Journal of Industrial Organisation*, **16,** 121-33.

Hull, J. C. (1995) *Introduction to Options and Futures Markets,* Prentice Hall, Englewood Cliffs.

Jing-Yuan, W. and Smeers, Y. (1999) "Spatial Oligopolistic Electricity Models with Cournot Generators and Regulated Transmission Prices", *Operations Research,* **47**(1)**,** 102-12.

Jones, T. (2000) *Electricity Demand Side Management*, Masters of Engineering Management thesis, Department of Engineering, University of Canterbury

Kahneman, D. and Tversky, A. (1979) "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica*, **47**(2)**,** 263-91.

Keers, G. (2000) "Taking the Corporate Risk out of Power Trading", *Petroleum Economist*, 67(3), 13-15

Kerr, A. L., Read, E. G. and Kaye, R. J. (1997) "Stochastic Dynamic Programming Applied to Medium-term Reservoir Management", Working Paper WP-97-03, Energy Modelling Research Group, University of Canterbury.

Keynes, J. M. (1938) "The Policy of Government Storage of Foodstuffs and Raw Materials", *Economic Journal*, **48,** 449-60.

Klemperer, P. D. and Meyer, M. A. (1989) "Supply Function Equilibria in Oligopoly under Uncertainty", *Econometrica*, **57**(6)**,** 1243-77.

Kriebel, K. W. and Hornstein, M. D. (1999) "Financing Merchant Power Plants", *International Financial Law Review*, July 1999, 30-34

Larsen, E. R. and Bunn, D. W. (1999) "Deregulation in Electricity: Understanding Strategic and Regulatory Risk", *Journal of the Operational Research Society*, **50,** 337-44.

Laughhunn, D. J., Payne, J. W. and Crum, R. (1980) "Managerial Risk Preferences for Below-Target Returns", *Management Science*, **26**(12), 1238-49.

Levy, H. and Markowitz, H. M. (1979) "Approximating Expected Utility by a Function of Mean and Variance", *American Economic Review*, **69**(3), 308-17.

Little, J. D. C. (1955) "The Use of Storage Water in a Hydroelectric System", *Operations Research*, **3,** 187-97.

Lowrey, C. (1997) "The Pool and Forward Contracts in the UK Electricity Supply Industry", *Energy Policy*, **25**(4), 413-23.

March, J. G. and Shapira, Z. (1987) "Managerial Perspectives on Risk and Risk Taking", *Management Science*, **33**(11), 1404-18.

Markowitz, H. M. (1959) *Portfolio Selection,* John Wiley and Sons, New York.

Mehra, R. and Prescott, E. (1985) "The Equity Premium: A Puzzle", *Journal of Monetary Economics*,, 141-64.

Mount, T. (2001) "Market Power and Price Volatility in Restructured Markets for Electricity", *Decision Support Systems*, **30,** 311-25.

Murtagh, B. A. and Murtagh, R. W. (1996) "A Model for Optimising Options and Futures", *International Transactions in Operations Research*, **2**(4), 399-404.

Neame, P., Philpott, A. B. and Pritchard, G. (2002) "Offer Stack Optimisation in Electricity Pool Markets", Working Paper, Electric Power Optimization Centre, University of Auckland.

Neumann, J. von. and Morgenstern, O. (1947) "The Notion of Utility", In *Theory of Games and Economic Behaviour*, Princeton University Press, Princeton, pp. 641.

Newbery, D. (1984) "Commodity Price Stabilization in Imperfect or Cartelized Markets", *Econometrica*, **52**(3), 563-78.

Newbery, D. (1998) "Competition, Contracts and Entry in the Electricity Spot Market", *RAND Journal of Economics*, **29**(4), 726-49.

Oren, S. (2001) "Integrating Real and Financial Options in Demand-side Electricity Contracts", *Decision Support Systems*, **30**, 279-88.

Pereira, M.V.F. and L.M.V.G. Pinto (1991) "Multi-stage stochastic optimisation applied to energy planning", *Mathematical Programming*, **52**, 359-375.

Pineau, P.-O. and Murto, P. (1999) "A Stochastic Dynamic Game Model of the Finnish Electricity Market", Working Paper, Helsinki University of Technology.

Powell, A. (1993) "Trading Forward in an Imperfect Market: The Case of Electricity in Britain", *The Economic Journal*, **103**, 444-53.

Pratt, J. W. (1964) "Risk Aversion in the Small and the Large", *Econometrica*, **32**(1), 122-36.

Ranatunga, R. A. S. K. (1995) *Risk Averse Operation of an Electricity Plant in and Electricity Market*, Master of Engineering thesis, School of Electrical Engineering, University of New South Wales

Read, E. G. (1984) "Deterministic Reservoir Operation - An Application of the Economic Principles", Working Paper, New Zealand Ministry of Energy.

Read, E. G. (1989) "A Dual Approach to Stochastic Dynamic Programming for Reservoir Release Scheduling", In *Dynamic Programming for Optimal Water Resources Systems Analysis* (Ed, Esogbue, A. O.), Prentice-Hall, Englewood Cliffs, N.J., pp. 361-72.

Rothwell, G. S. (2000) "The Risk of Early Retirement of U.S. Nuclear power Plants under Electricity Deregulation and CO2 Emission Reductions", *The Energy Journal*, **21**(3), 61-87.

Sarris, A. H. (1984) "Speculative Storage, Futures Markets and the Stability of Commodity Prices", *Economic Inquiry*, **22**, 80-97.

Schmalensee, R. and Golub, B. (1985) "Estimating Effective Concentration in Deregulated Wholesale Electricity Markets", *RAND Journal of Economics*, **15**(1), 12-26.

Scott, T. J. (1998) *Hydro Reservoir Management for an Electricity Market with Long-term Contracts*, Ph.D. thesis, Department of Management, University of Canterbury

Scott, T. J. and Read, E. G. (1996) "Modelling Hydro Reservoir Operation in a Deregulated Electricity Sector", *International Transactions in Operations Research*, **3**(3/4), 243-53.

Shapley, L. (1953) "Stochastic Games", *Proceedings of the National Academy of Science*, **39**, 1095-100.

Smeers, Y. (1997) "Computable Equilibrium Models and the Restructuring of the European Electricity and Gas Markets", *The Energy Journal*, **18**(4), 1-31.

Smeers, Y. and Boucher, J. (2001) "Alternative Models of Restructured Electricity Systems, Part 1: No Market Power", *Operations Research*, **49**(6), 821-38.

Smeers, Y. and de Wolf, D. (1997) "A Stochastic Version of a Stackelberg-Nash-Cournot Equilibrium Model", *Management Science*, **43**(2), 190-97.

Sortino, F. A. and Meer, R. v. d. (1991) "Downside Risk", *The Journal of Portfolio Management*, **17**(4), 27-31.

Tirole, J. (1988) *The Theory of Industrial Organisation,* Massachusetts Institute of Technology Press, Cambridge, Massachusetts.

Turnovsky, S. (1983) "The Determination of Spot and Futures Prices with Storable Commodities", *Econometrica*, **51**(5), 1363-87.

Wambach, A. (1999) "Bertrand Competition under Cost Uncertainty", *International Journal of Industrial Organisation*, **17**, 941-51.

Ward, S. (1997) "Managing Risk - A Key Task for Management", *OR Insight*, **10**(2).

Weber, E. U. and Milliman, R. A. (1997) "Perceived Risk Attitudes: Relating Risk Perception to Risky Choice", *Management Science*, **43**(2), 123-44.

Weyant, J. P. and Hill, J. (1998) "The Costs of the Kyoto Protocol: a Multi-Model Evaluation", *The Energy Journal*, **Special Issue:The Costs of the Kyoto Protocol,** vii-xliv.

Williams, J. (1986) *The Economic Function of Futures Markets,* Cambridge University Press, Cambridge.

Williams, J. C. and Wright, B. D. (1991) *Storage and Commodity Markets,* Cambridge University Press, Cambridge.

Wolak, F. A. (1999) "An Empirical Analysis of the Impact of Hedge Contracts on Bidding Behaviour in a Competitive Electricity Market", *Fourth Annual POWER Research Conference*, Berkeley, CA, 50

Wolak, F. A. and Patrick, R. H. (1996) "The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Market", Working Paper PWP-047, University of California Energy Institute.

Wolfram, C. D. (1999) "Measuring Duopoly Power in the British Electricity Spot Market", *American Economic Review*, **89**(4), 805-26.

Yakowitz, S. (1973) "A Stochastic Model for Daily River flows in an Arid Region", *Water Resources Research*, **9**(5), 1271-85.

Yakowitz, S. (1982) "Dynamic Programming Applications in Water Resources", *Water Resources Research*, **18**(4), 673-96.

Yang, M. (1995) *Dual Dynamic Programming for Reservoir Management with Correlated Inflows*, PhD thesis, Department of Management, University of Canterbury

Yeh, W. W.-G. (1985) "Reservoir Management and Operations Models: A State-of-the-Art Review", *Water Resources Research*, **21**(12), 1797-818.

A RISK

A.1 Introduction

This Appendix will outline the existing literature on decision making under risk. Given the breadth of treatment risk is given in the literature, we only intend to highlight the fundamental aspects, and the relevant papers, since this thesis is not concerned with the most accurate representation of decision making uncertainty in electricity markets.

A.2 Decision Making Under Risk

The topic of risk has received a considerable amount of attention in the literature. The majority of the discussions surrounding risk are found in the decision theory literature, which should not be surprising, given that the majority of decisions in managerial circles are made in the context of uncertainty. The notion of risk to most people involves negative outcomes and the idea that those outcomes are uncertain. However, these loose definitions are insufficient when it comes to models of decision making – the desire for more concrete ideas of defining and representing attitudes to risk is evident in the burgeoning literature surrounding these topics. A complete review of this literature would be near impossible, as the range of theoretical attempts to represent different personal approaches to risk is compounded by the plethora of empirical studies of how closely theory matches the reality of managerial perspectives of risk. The reader is referred to March and Shapira (1987) for an excellent discussion and literature review of comparisons of the basic tenets of decision theory with empirical analyses of actual managerial perspectives.

However, it is incumbent upon the current study to examine the basic structure of what might loosely be termed "risk analysis":

- The definition of risk
- Decision makers' attitudes to risk, and
- The effect of this attitude on decision making.

These aspects of risk analysis are not mutually exclusive. Given that it is the purpose of this thesis to understand the complex interactions involved in decision making in an uncertain market, this discussion must go some way to understanding the complexity of risk in order to understand the limitations of the insights that will later be drawn.

A.2.1 Frameworks for Defining Risk

Specific definitions of risk are as varied as the methods of dealing with it. This is not helped in the least by the controversy surrounding risk. Fischhoff, Watson and Hope (1984) comments:

> *"The choice of definition [of risk] can affect the outcome of policy debates, the allocation of resources among safety measures and the distribution of political power in society.....No definition [can be] advanced as the correct one, because there is no one definition that is suitable for all problems. Rather, the choice of definition is a political one, expressing someone's views regarding the importance of different adverse effects in a particular situation"*

The general conception in economic analyses is that risk represents the possibility that the outcome of a gamble will turn out worse than expected Cozzolino (1979). The Oxford Dictionary defines risk as "the chance of injury or loss". Thus the concepts of uncertainty, probability and expected value are inextricable in understanding risk.

Risk and uncertainty are by no means equivalent. The traditional treatment of these two concepts, summarised by Fishburn (1984), is that uncertainty refers to probabilities between 0 and 1, and distributions or outcomes with such probabilities, while risk is viewed in the conventional manner of bad outcomes with positive probabilities. However, in an excellent literature survey on the subject of risk, Ranatunga (1995) points out that many modern authors now agree that uncertainty is a concept that is not measurable; risk is that part of uncertainty that is measurable.

The above definition of risk also seems to imply that risk is two-dimensional – both the probability of the negative outcome, and the magnitude of the outcome, work simultaneously to determine risk. The natural corollary here is that risk is increasing in the probability of a negative outcome, given certain magnitude, and increasing in magnitude (in the negative sense) for a given probability.

Of course, a 'gamble', in the wider sense of risk, could involve the possibility of catastrophic events (for example nuclear accidents), or more generally, unrepeatable and/or immeasurable events. Including these outcomes in a framework of measuring risk is further complicated by the fact that their occurrence tends to be beyond an individual's ability to assign a probability. Hence restricting our discussion to areas of uncertainty that are both measurable in terms of value, and able to have probabilities assigned, will be helpful.

Classical decision theory treats risk in such a way, reflecting the variation in the distribution of possible outcomes, their likelihoods and their subjective values March and Shapira (1987). Decision theorists use two prominent frameworks to achieve this:

1. The use of numerical utility functions to represent risk preference, or,

2. Directly measuring the statistical variance of the distribution of outcomes, and usually trading off levels of risk with the expected return of the alternative.

Utility functions, generally speaking, circumvent the direct measurement of risk itself; by placing a value on a decision makers "feeling" or level of satisfaction of a certain monetary outcome, a distribution of outcomes is transformed into a distribution of utilities Weber and Milliman (1997). The decision makers attitude towards the distribution of outcomes is then reflected by manipulating the exact form of the utility function (see Section A.2.2).

On the other hand, risk-return models, as reported by Bell (1995), are often more "intuitively satisfying" than its utility counterpart, since it explicitly provides a trade-off between the expected return of a gamble with its associated risk. However, the use of the variance of possible outcomes to reflect the riskiness of a venture is the subject of controversy itself. In such a formulation, the more variable the distribution, the more risky the gamble is considered. This in part reflects the fact that the more uncertain a decision maker is about what return will be made, the more risky the decision is considered.

Dyer and Sarin (1982) first introduced the idea of relative risk aversion as a conception of two cognitive representations, namely risk preference from risk perception. Risk preference could be loosely defined as "how much risk do I want", while risk perception is an assessment by an individual as to how risky a particular gamble is, two factors "confounded in the expected utility framework". Weber and Milliman (1997) claim that using a single utility function for risk may cloud these two issues, especially when trying to explain changes in choice, and in particular, sub-optimal choice:

> *"For purposes of decision aiding or remediation of suboptimal choice behaviour it is crucial to know which of these mechanisms determines observed changes in choice. If changes in risk perception are the driving force, then effective remediation should target cognitive processes, with information aimed at more realistic risk perception. If changes in risk preference are the driving force, then intervention needs to target people's emotional responses"*

Fishburn (1984) proposes an axiomatic treatment of risk that jointly considers the probability of loss, and the distribution of losses. In this way, Fishburn's model of risk could be considers to fit in with the ideas of semi-variance and downside risk (Section A.3.3), as he sets out to measure individuals preferences towards distributions that contain outcomes below a target, which are considered to be "undesirable" or "risky". Fishburn's axioms for risk measurement are based upon the statement that one decision alternative is "at least as risky as" another.

A.2.2 Expected Utility Hypothesis

Utility theory has become central to an understanding of individuals attitude to risk. The notion of utility dates back as far as the 18^{th} century, but the major advance in utility functions, which map the value of the outcome, x, to the utility, $U(x)$, was made by John Von Neumann and Oskar Morgenstern in 1947. von Neumann and Morgenstern acted in response to Bernoulli's observation that individuals refused to take a gamble even though the expectation of the gamble was infinite. Bernoulli denounced the use of mathematical expectation as being inaccurate in explaining the psychological behaviour of such individuals, and coined the term "moral expectation" to represent the gamblers attitude. In essence, von Neumann and Morgenstern wanted to define a numerical transformation of the monetary outcomes so that mathematical expectation could be used. The overall utility of a gamble would then be the expected utility of its outcomes. These transformations were called utility functions Neumann and Morgenstern (1947).

Utility is commonly conceived as reflecting an individuals feeling about a particular (usually monetary) outcome, so it should be clear how useful it is to modelling risk. Von Neumann and Morgenstern formalised the use of utility by defining axioms (which is essentially a set of rational choice assumptions) for utility functions. While the axioms, which essentially defined the cardinality of the functions, will not be repeated here, they concluded that:

> *"The fact that a numerical utility – with a formula amounting to the use of mathematical expectations – can be built upon [these axioms], seems to indicate this: We have practically defined numerical utility as being that thing for which the calculus of mathematical expectations is legitimate." Neumann and Morgenstern (1947))*

Thus, for two alternative gambles $\tilde{x}$ and $\tilde{y}$ with utilities $U(\tilde{x})$ and $U(\tilde{y})$ and mathematical expectation over the underlying probability distribution *E(.)*, we can say that if

$$E(U[\tilde{x}]) > E(U[\tilde{y}])$$

then $\tilde{x}$ is preferred to $\tilde{y}$. Any utility function satisfying the axiomatic treatment of von Neumann and Morgenstern is termed a von Neumann and Morgenstern (vNM) utility function. The compelling nature of these authors' axioms has been a major factor in vNM utility functions surviving decades of criticism Bell (1995).

A.2.3 Risk Aversion

Inherent in any attitude to decision making is the notion of preference, i.e., given two options, I am able to say which I prefer or value higher, based on some criteria. The earliest idea of risk aversion, pioneered (among others) by Arrow (1971), reflects this choice by defining a risk averse individual as one who will prefer a certain outcome over a gamble that has the same expected value as the certain one(an actuarially fair bet). Extending the utility theory outlined above, Arrow went further to define two fundamental properties of rational and risk averse utility functions:

1. **More is Better**: Wealth is always desirable, or, more formally, marginal utility is always positive, i.e., $U'(x) > 0$

2. **Decreasing Marginal Utility**: Marginal utility is inversely proportional to wealth, i.e., as wealth increases, $U'(x)$ decreases. Formally, $U''(x) < 0$.

This situation is represented geometrically below. An individual is offered a lottery L involving two outcomes, x and y, of equal probability. If either x or y were received with certainty, the investor would have utility $U(x)$ or $U(y)$ respectively, according to the utility function $U(w)$ (Figure A.1 (a)). The expected utility to the investor is the weighted average of these two individual utilities (each with probability $p = 0.5$), depicted in Figure A.1 (b) where the weighted average of the outcomes, *E[L]*, intersects a straight line between the two points. However, this is the same utility as the investor would have of a certain amount z, which is a lower monetary value than the expected value of the lottery. We can say therefore the investor is indifferent between the certain amount z and a lottery with expected value *E[L]*, where $z < E[L]$.

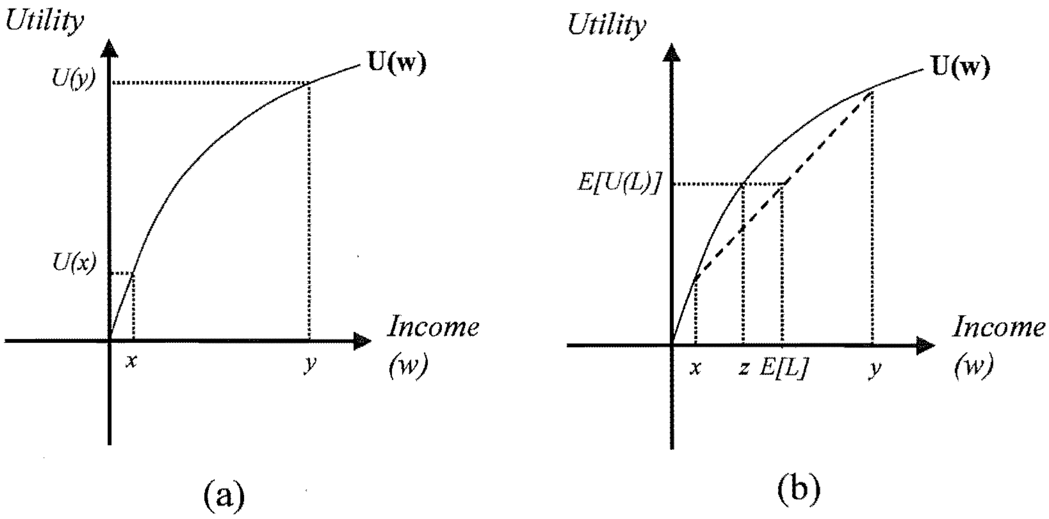


Figure A.1 Risk averse utility functions

These conditions on risk averse utility amount to a definition of concave utility functions given a distribution of wealth. The natural extension to this are the concepts of risk neutrality and risk seeking behaviour, where individuals are indifferent between the gamble and its certainty equivalent in the former case, and prefer the gamble (convex utility) in the latter.

These functions, as von Neumann and Morgenstern intended, are subjected to mathematical expectation, so that an individual maximises his expected utility rather than expected monetary wealth where utility is weighted in exactly the same fashion as the mathematical treatment of the latter. This is known as the expected utility hypothesis.

Earlier work by Arrow (1971) and Pratt (1964) contended that human beings were all basically risk averse. Intending to formalise this belief, Arrow and Pratt searched for measures of risk aversion. It would be tempting to use the rate of change in marginal utility, $U''(x)$ for condition 2 above, implying that curvature of the utility function indicated the degree of risk aversion. However, Arrow points out that this would violate the requirement that the preference ordering implied by a utility function be invariant under linear transformations. Multiplying $U(x)$ by a constant does not imply any behavioural changes – the preference order remains the same. However, $U''(x)$ does change under constant multiplication of $U(x)$, therefore the numerical value of $U''(x)$ has no significance. Hence we have the important result that $U''(x)$ implies risk aversion, but its numerical value does not help us compare degrees of risk aversion Arrow (1971).

Instead, Arrow and Pratt defined two measures of investors risk aversion, absolute risk aversion denoted $R_A(x)$ and relative risk aversion, denoted $R_R(x)$. These remain invariant under linear transformation and are defined as:

$$R_A = -\frac{U''(x)}{U'(x)}$$

$$R_R = -\frac{xU''(x)}{U'(x)}$$

The values of $R_A(x)$ and $R_R(x)$ change as wealth (x) changes. The second definition is analogous to the elasticity of marginal utility with respect to the investor's wealth. Francis and Archer (1979) provides a good description of how these measures should be interpreted.

Pratt (1964) gave a useful intuitive interpretation of the risk aversion measures. Consider, as in the case above, an individual who is indifferent between an uncertain distribution of income $\tilde{Y}$, and alternatively, a certain amount Y_0. A risk averse individual will choose Y_0 to be less that $E(\tilde{Y})$, in accordance with the general utility function shown in Figure A.1 above. Pratt interpreted the difference between $E(\tilde{Y})$and Y_0 as the "risk premium". Pratt then shows that

$$\pi(\tilde{Y}) = \frac{1}{2}\sigma^2 R_A(\tilde{Y}) + o(\sigma^2)$$

where σ^2 is the variance of the gamble $\tilde{Y}$, and $o(.)$ can be interpreted as "terms of smaller order than". $\pi(\tilde{Y})$ is the risk premium such that the investor would be indifferent between receiving the certain amount $Y_0 = E(\tilde{Y}) + \pi(\tilde{Y})$ and receiving the uncertain income stream $\tilde{Y}$. Hence the more absolutely risk averse an individual is, the greater the risk premium and therefore certainty equivalent required.

As Arrow later pointed out, risk aversion is not global: if all humans were completely risk averse, why does organised gambling succeed? Friedman and Savage (1948) suggested that humans were possibly risk averse to some risk and not to others – in particular, the larger the amounts involved, the stronger influence risk aversion will have.

Other criticisms of the global risk-aversion inference of concave utility functions include the idea of Friedman and Savage that individuals risk attitude is a mixture over wealth, states and other situational contexts. Kahneman and Tversky (1979) narrowed this further by suggesting that individuals often appear to be risk averse when making net gains or profits, and risk-seeking when making losses. This effect has been empirically supported by Laughhunn, Payne and Crum (1980).

A.3 Risk-Return Models

Most risk-return frameworks propose that the overall choice behaviour involves both the expected value of the alternative, and its riskiness as defined by the variability of the distribution of outcomes. Bell (1995) comments that this explicit separation of risk and return as primitives is often more representative of management attitudes, something not modelled *per se* under the expected utility hypothesis. He suggests:

> *"...informal discussion of alternatives by decision makers often includes statements such as 'alternative A is more attractive than alternative B, but is too risky,' suggesting that decisions are thought of, at an intuitive level at least, as a trade-off between the risk inherent in the alternative, and their levels of 'return' (their attractiveness were it not for the risk)."*

As stated previously, the alternative that has higher variance is more risky (and less attractive to the risk-averter). As was the case under the more-is-better utility framework, the alternative which has a higher expected value is more attractive. Thus, expected value has a positive correlation, and variance a negative correlation with the overall attractiveness of an alternative. This forms the basic risk-return trade-off faced by the risk averse decision maker – once all the risk-return combinations have been calculated, the decision maker selects the one that is most attractive given his risk attitude.

From the basic risk-averse choice outlined in the certainty-equivalent problem, it also follows that individuals must be compensated for increased variance by increased expected return. Thus the greater the return on a gamble, the greater the risk that should be involved (March and Shapira (1987)). Stated another way, risk-averters will sacrifice expected return on investments in order to achieve lower levels of risk.

A.3.1 Portfolio Theory

Possibly the foremost treatment of risk-return trade-offs was by Harry Markowitz, in his development of portfolio theory. Markowitz was among other authors that doubted the intuitive value of expected utility theory:

> *"[The utility] approach will probably have less immediate meaning and intuitive appeal for him than an analysis in which the investor is shown combinations of "risk" and "return" and is then asked to pick carefully the combination that best suits his needs. Choosing a combination of risk and return is a more natural procedure than expressing attitudes towards risk in terms of a utility function and then leaving the choice to a machine"* Levy and Markowitz (1979)

A portfolio is a collection investments, each with their own risk and return. Rather than evaluateing each investment in its own right, portfolio management looks at the interactions of a number of investments owned by the same investor. The value of a given portfolio is defined as a weighted average of the returns of the individual assets contained within it. These weights are usually simply the proportion of the total monetary investment allocated to that asset. One dollar invested in asset *a* will yield the investor a payout of $(1+\tilde{r}_a)$ dollars, where $\tilde{r}_a$ is the uncertain rate of return on asset *a*. So the rate of return of the portfolio is:

$$\tilde{r}_P = q_1\tilde{r}_1 + q_2\tilde{r}_2 + q_3\tilde{r}_3 + \ldots q_A\tilde{r}_A$$

where q_a is the proportion of total initial wealth invested in that asset. The variance of the return on the portfolio can be calculated in a similar way, so that the variance, $V(\tilde{r}_P)$ is a function of the individual variances of the assets, and the covariance between the assets' returns.

The original precepts of Portfolio Management display all portfolios of investments in the 2-dimensional space of risk and return, according to their given combination of the two primitives. An feasible combination set for a two-asset portfolio shows the resulting portfolio risk and return as the weights are varied between 0 and 1 for each asset. The familiar quadratic-shaped feasible set for the two-asset portfolio is given in Figure A.2. Two portfolios, A and B, represent different combinations of these assets.

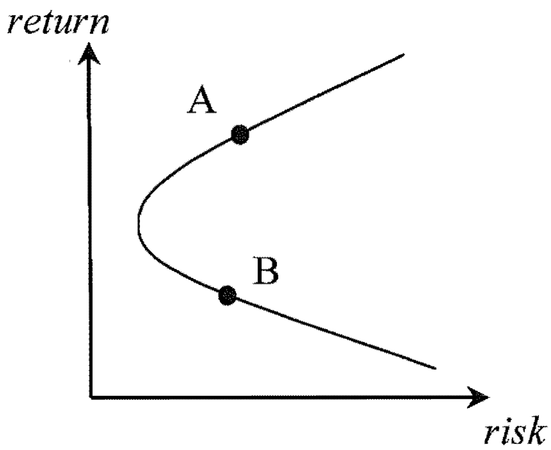


Figure A.2 Risk and return of feasible combinations of two assets

Any portfolio has an "efficient frontier", a representation of the 'best' combinations of investments in risk-return space. For an investor obeying the assumptions of vNM risk averse utility functions (i.e., concave utility in wealth), this is clearly the upper half of the hyperbola in Figure A.3(a), since each combination of risk and return dominates those on the lower half (higher return for the same level of risk). Where exactly an investor positions himself on the frontier depends on his attitude to risk, or utility function. In fact, the optimal portfolio choice is found where the investors utility indifference curve is just tangent to the efficient frontier (Figure A.3(b)).

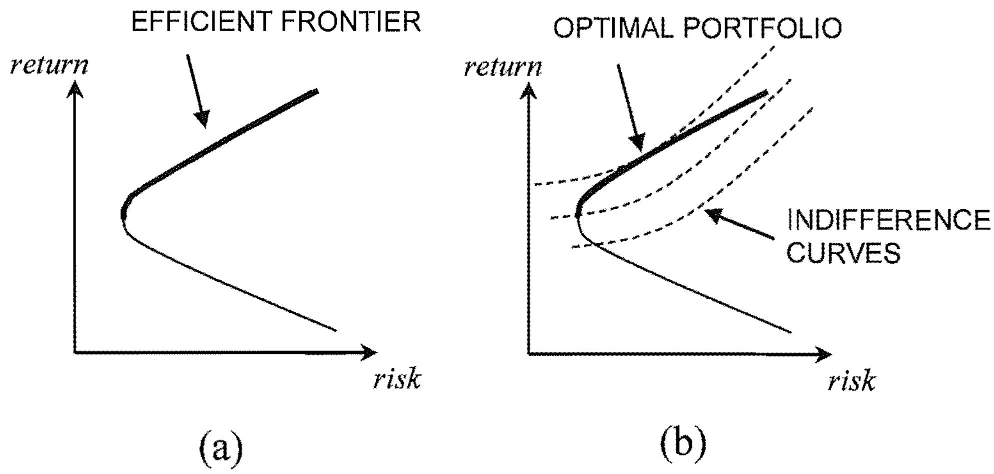


Figure A.3 The efficient frontier and optimal portfolio selection

Portfolio optimisation techniques go beyond this processing stage by choosing which assets should be included in the portfolio. Markowitz diversification is defined as combining assets that are less than positively perfectly correlated so as to reduce the overall variance of the portfolio without compromising the expected return. Other optimisation analyses look at multiperiod portfolios, using the operations research technique of dynamic programming Francis and Archer (1979).

A good summary of modern portfolio theory can be found in Elton and Gruber (1995).

A.3.2 Mean-Variance

While the risk-return and expected utility treatments of risk differ, each can be argued to be "rational conceptualisations" of decision making under risk. It is not unreasonable therefore to expect that they should exhibit some correspondence with each other (Weber and Milliman (1997)). Authors that have attempted to find the compatibility of these models, which essentially amounts to finding classes of utility functions which represent risk-return trade-offs, include Bell (1995), Jia and Dyer (1995) and Sarin and Weber (1993).

One area of compatibility was developed using a mean-variance model. The general form of mean variance functions is to define utility of some uncertain outcome $\tilde{x}$, $U(\tilde{x})$ such that:

$$U(\tilde{x}) = \mu_x - \lambda\left(\sigma_x^2\right)$$

Where μ_x is the mean and σ_x^2 the variance of the outcome distribution (thus risk is measured by the variance of the gamble). λ then represents the degree of risk aversion, and could be interpreted as how much the individual penalises the expected outcome by to represent their risk preference and the perceived riskiness of the gamble.

Levy and Markowitz (1979) manipulated the relationship between risk and variance to show that certain classes of mean-variance frameworks could be approximated by quadratic utility functions, since both quadratic utility and variance were second-order polynomials in wealth.

Consider the quadratic von Neumann Morgenstern utility function:

$$U(x) = ax - \frac{1}{2}x^2$$

For $a > x$, this function satisfies the general requirements of risk averse utility, since $U'(x) = a - x$ and $U''(x) = -1$ are both strictly negative (see below for a commentary on this restriction).

Under the expect utility hypothesis, for some uncertain income $\tilde{x}$:

$$\begin{aligned} E\left[U(\tilde{x})\right] &= E\left[a\tilde{x} - \frac{1}{2}\tilde{x}^2\right] \\ &= aE\left[\tilde{x}\right] - \frac{1}{2}E\left[\tilde{x}^2\right] \end{aligned}$$

Let μ equal the expected value and σ^2 the variance of the income distribution. Then

$$\mu = E[\tilde{x}]$$

and

$$\begin{aligned}\sigma^2 &= E\left[\tilde{x}^2\right] - \left(E[\tilde{x}]\right)^2 \\ &= E\left[\tilde{x}^2\right] - \mu^2\end{aligned}$$

and we can thus say that

$$E\left[\tilde{x}^2\right] = \sigma^2 + \mu^2$$

Therefore, expected utility can be written

$$E[U(\tilde{x})] = a\mu - \frac{1}{2}\left(\sigma^2 + \mu^2\right)$$

i.e., as a combination of the expected value and variance (risk) of the income stream.

The problem with quadratic utility is quite clear (Figure A.4), and was strongly criticised by Arrow (1971) and others for a lack of sense in its global properties, namely:

- Utility is not increasing in wealth everywhere
- Risk aversion is not globally evident, since $\dfrac{1}{U'(x)}$ is unbounded
- Risk aversion is, in the region greater than (less than for the positive quadratic) the apex, increasing in wealth

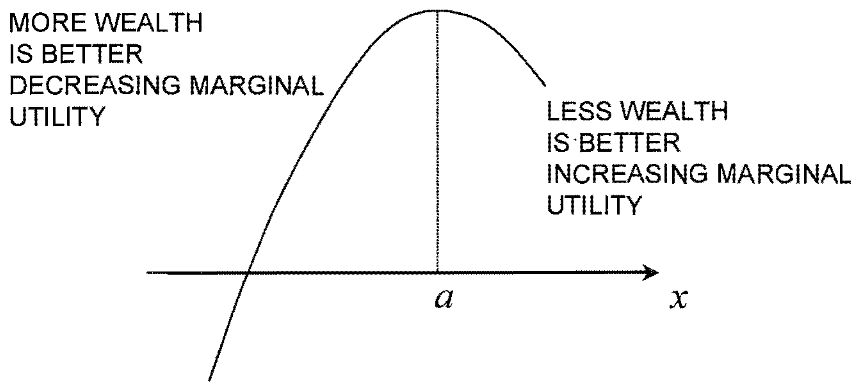


Figure A.4 Unrealistic quadratic utility

This absurdity of quadratic utility was largely dealt with by authors who supported its use developing restrictive assumptions, mainly that it is only an accurate approximation to true risk-averse preferences over the region where the function is increasing. Markowitz (1959) gives a good exposition of the criticisms, and defences and clarifications of the use of quadratic utility functions.

Additionally, the use of variance has been criticised on two main points:

- Variance includes both positive and negative outcomes. Thus an increase in the frequency of favourable outcomes would be interpreted as a more risky option, which for some decisions is counter-intuitive. This point drove the development of "downside risk", which generally measures that part of the distribution below the mean. This is not to say that positive outcomes should be ignored (as in the semi-variance or downside risk approach in Section A.3.3). Fishburn (1984) points out that the presence of favourable outcomes in a distribution with unfavourable outcomes will still have an effect on the overall riskiness of the distribution and an individual's preference towards it.

- Variance does not totally capture the shape of a distribution. One could easily construct an example where two distributions have identical variance, but have very different shapes. For example, one option may have outcomes distributed evenly around the mean, and a second may have a high probability of small positive outcomes, and a very small probability of a very damaging outcome.

Many decision makers would not consider these options equally risky. This is especially of relevance to decisions which may have catastrophic outcomes with a very small associated probability, and thus their effect is "swamped" in the calculation of variance.

The benefits of mean-variance functions are that our attention to utility can be restricted to simple functions of two easily-obtainable moments of the underlying distribution, greatly simplifying analysis (Guthrie (1998)). Markowitz (1959) points out that variance has significant advantages in computational cost, convenience and familiarity to decision makers with a passing acquaintance with modern statistics. Many would contend that these benefits compensate for the restrictive assumptions and/or the lack of explanatory power of the quadratic and mean-variance frameworks. This debate still continues.

A.3.3 Downside Risk

In order to deal with the observation that upside opportunities are treated in the same fashion as downside risks in the mean-variance framework it would be helpful to consider only those parts of the distribution that do constitute a loss, or below target return for the individual. This would more accurately reflect the definition of risk given in Section A.2.1.

Downside risk may be especially relevant to investment decisions. Often investment decisions require a minimum rate of return to ensure the viability of a project, or that a surplus will be generated Sortino and Meer (1991). Chapter 2 details an example from the electricity industry, where high priced thermal plant is only concerned with that part of the price distribution that will render it undispatched, i.e., where price falls below marginal cost.

Fishburn and Markowitz was among the early authors who developed a vN-M utility function for an individual who is averse to downside risk. The general model as quoted in Ranatunga (1995) is:

$$U(x) = \begin{cases} x & x \geq h \\ x - b(h-x)^a & x < h \end{cases}$$

for some *a*, $b > 0$, and a target or expected wealth *h*. This is similar in some respects to the general quadratic utility model, i.e., for $a = 2$. This amounts to the utility depicted in Figure A.5, where a decision maker exhibits quadratic utility below a target, and linear (risk neutral) utility above the target (in the case of Figure , the target is *h*).

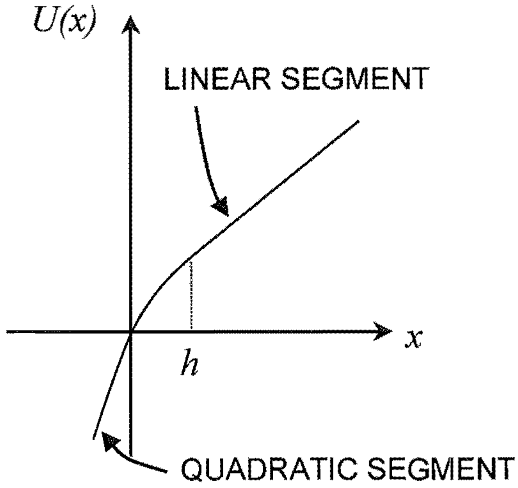


Figure A.5 Use of a piecewise utility to reflect downside risk

Markowitz (1959) took the more direct approach of replacing variance with "semi-variance" in a mean-variance framework. For some risky income, *r*, Semi-variance is defined as:

$$S = E\left(r-\right)^2$$

where

$$r- = \begin{cases} r & r \leq 0 \\ 0 & r > 0 \end{cases}$$

assuming that the expected value or target outcome is 0. It is noted that measuring risk by semi-variance is equivalent (in terms of preferences, but not numerical value) to variance when the distribution is symmetric.

Both Sortino and Meer (1991) and Markowitz (1959) highlight the computational disadvantages of downside risk, especially within a risk-return portfolio framework with continuous underlying distributions. The estimation of downside risk amounts to the calculation of an integral, before any portfolio optimisation can begin.

A.4 Prospect Theory

There are many complications to the two broad frameworks outlined above, and it is not a difficult task to cite examples of decision making under risk that appear to violate the assumptions or contravene the results of the risk-return or expected utility hypotheses.

One are of major contribution has been the 'prospect theory' of Kahneman and Tversky (1979). A 'prospect' is defined as a "contract that yields an outcome x_i with probability p_i, such that $p_1 + p_2 + p_3 + ... + p_n = 1$". The authors presented a series of relatively complex decision problems, for which the preferences were found empirically, that seemed to contravene the axioms of expected utility theory. They then went on to define the particular deviations from the expected utility hypothesis:

The certainty effect: In EU theory, the utilities of outcomes are weighted by their probabilities. However, Kahneman and Tversky showed that individuals tend to overweight certain outcomes relative to their probable (high probability) counterparts

The reflection effect: This builds on earlier work that postulated that individuals estimate the risk of a gamble by the deviations of the outcomes from a reference point, rather than the final wealth. The reflection effect shows that preferences between loss prospects were completely opposite to preferences between gain prospects. In other words, individuals tended to be risk-seeking in the loss domain and risk-averse in the gain domain.

The isolation effect: While EU theory ascertains that individuals develop preferences based on the entire distribution of outcomes, Kahneman and Tversky found that individuals often decompose prospects into common and distinctive components, and often discarded the common, leaving final preferences to be made based on the distinctive.

Kahneman and Tversky's prospect theory was in essence a method for pre-processing gambles into their simplest and most descriptive form, and then evaluating preferences.

B ANALYSIS OF PROFIT CONCAVITY

Introduction

Chapter 11 developed a set of conditions that would describe a market equilibrium where both firms would attempt to manipulate both the variance and covariance terms, in the contract demand equation, to their advantage. The second order condition to problem 11.2 yielded a complicated expression:

$$0 > -2b - c_i + b^2 \lambda k_i \left(G^r - E[G] - \left(K - G^r\right)\left(2 - 2\theta_i^r\right)\right) \tag{11.23}$$

It is difficult to derive an intuitively appealing rearrangement of 11.16, in order to describe the cobinatioOns of parameters which will lead to the conditions, outlined in Chapter 11, describing a profit maximum. Chapter 12 outlined numerical results for a range of "sensible" parameter values, and solutions were found to satisfy the second order condition. Since the direction of this thesis was to illustrate that an effect existed, i.e., that firms may find equilibria in the market that improve profit by increasing the risk faced by consumers, we believed that traversing the somewhat difficult mathematical territory, where conditions on the parameters were explored, would not add anything to the general thrust of the thesis. Simple numerical examples were presented in Chapter 11 that illustrated that there can be situations where profit maximising equilibria cannot be found, and thus Cournot reaction functions cannot be defined, whereas the analytical investigation is presented here.

In order to simplify our discussion of (11.19), we will initially assume that the firm is a monopolist (and hence G^r is the generation of the firm in state r) so as not to complicate our understanding with the strategic interaction between the firms. In the monopoly case, first order condition (11.15) is the same, except the variables of other firms' generation are no longer relevant. Hence incentives still exist to "spread" prices around the zero-contract monopolist solution.

Monopoly Analysis

For this case, equation (11.19) can be rearranged in numerous different ways in order to find an intuitive explanation. Given the first-order incentives to spread prices (and, by implication, generation), a region of convex profit will occur when such spreading is increasingly attractive relative to the sacrificed short-run spot profits. In these regions, the firm faces to spread prices even further, and the value of the first-order derivative moves further away from zero (and further away from a solution). If a solution cannot be found within the concave-profit region, no profit maximising solution exists to the monopolist's problem.

Given this interaction between first-order solutions, price spreading and profit concavity, it is helpful to rearrange the second-order condition to be a restriction on the magnitude of the generation "spread"[86]:

$$G^r - E\left[G^{-r}\right] < \frac{2}{3}\left(K - E\left[G^{-r}\right]\right) + \frac{2b+c}{3b^2\lambda K\left(1-\theta^r\right)} \tag{11.24}$$

where

$$E\left[G^{-r}\right] = \left(\frac{\sum_{t\neq r}\theta^t G^t}{\left(1-\theta^r\right)}\right)$$

Whether or not condition **Error! Reference source not found.** is restrictive on the firms desire to spread generation depends on the values of the right-hand side of the equation, the exact value of which is a numerical question. However, it is useful to note the effect of the critical parameters on how restrictive the condition is.

The fractional term on the right-hand side is strictly positive for any demand curve that is not perfectly elastic. The magnitude of the fractional term, and the sign of the bracketed term, are dependent on the level of contracts signed in equilibrium, although the effect of contracts on each is inversely related. By letting contracts tend to their extremes of zero and infinity, it can be seen that the right hand side is infinite in both cases, and thus not restrictive on the spread of prices from the mean. Hence, this condition will be at its most restrictive at some level of contracts between these two extremes, although the exact value remains a numerical question. Since we would expect that a firm is most likely to be under-contracted in equilibrium, the bracketed term will tighten the restriction on the ability of the firm to push generation above the mean.

An increase in the consumers' risk aversion will reduce the fractional term, and thus increase the degree of restriction of condition 11.24 on generation. This appears intuitive, given the discussion above which outlined the nature of the first-order tradeoff between contract profit and spot profit. As consumers become more risk-averse, risk premiums increase and the profitability of increasing the consumers risk through destabilisation also increases. Hence total profit will become convex at a lower degree of spreading.

The point at which the first-order solution, for a given state, becomes unstable is also dependent on the probability of that state occurring. For state probabilities close to 1, the fractional term of 11.24 is very large, and there is little restriction on the firm's desire to push generation high in low-cost states. However, low probability states appear to be the most restrictive (all other factors held equal), since values of θ_i^r close to zero minimise the fractional term. This seems counterintuitive, since the chance of finding a solution to the whole LRE system is thus largely driven by states that may hardly ever occur, and it also suggests that the attractiveness of destabilising strategies increases as the chance of the state occurring decreases.

[86] Equivalently, the condition can be rearranged to form a limit on the spread of generation above the contract level. Both types of equations indicate that there is an upper bound on generation, above which profit becomes convex.

However, the interpretation of the effect of probability is limited because of a simplification made earlier. In the development of first-order condition (11.15), it was noted that both the effect of a state's generation on contract revenue, via the spot price distribution, and the short-run spot profit, were weighted by the probability of the state occurring, and hence the probability could be removed from the first-order equation. So while it is true that the point at which potential increases in contract revenue begin to dominate sacrificed spot profits is low for small probability states, the incentives to destabilise at all in these states is also very low, given that such states have a very small overall impact on the consumers' spot price distribution[87].

Assuming that the right-hand side of 11.24 is at least positive, it is clear that the condition is most likely to be binding in those states in which generation exceeds the mean, i.e., states in which marginal costs are lower than average. We know that in these states, the firm faces first-order destabilising incentives to push generation higher than the short-run profit maximising output level, and thus closer to the bound described above (if they don't violate it already). In high-cost states, not only are generation levels low, destabilising incentives will lead the firm to decrease generation, thus moving it further away from the bound implied by the concavity condition.

Oligopoly Analysis

If there are combinations of these parameters which do not satisfy the second order condition for a monopolist, then it seems equally likely that the same will be true of multiple firms facing similar incentives in the contract market. In this case, G^r represents the total output of all firms, which, under short-run profit maximisation, is higher in equilibrium than the corresponding optimal monopoly output. However, total contracts, and the average output over all other states, would also be higher, so it is not immediately clear whether condition 11.24 would be more or less restrictive. However, the same basic intuition applies: Given the level of the firm's rivals' output, if the

[87] The belief that small probability states have little effect on spot price distributions is quite reasonable (even though their effect on perceived risk may be high). However, the implied assumption in the first order condition, that the firm is least concerned about spot losses in the lowest probability states, is questionable. While this is a direct result of the firm's objective of maximising expected profit, it seems unlikely that, in the few years in which the cost state is

solution to the first-order condition does not lie within the range of positive generation levels implied by 11.24, no profit maximising solution exists for that state.

This has important implications for the search for an equilibrium in the spot market, an issue we will now consider in depth.

13.4.1 Existence of a Spot Market Equilibrium

In the two-firm duopoly, an equilibrium exists if the profit-maximising solution to the first-order condition, given the other firm's output, induces the other firm to produce that output in response. Hence any sensible discussion of reaction functions is predicated on the assumption that a profit maximising solution to the first-order condition can be found, given a certain level of the other firm's output. Given that we have just shown that this may not be possible, particularly for high levels of generation corresponding to low-cost states, it is conceivable that there are regions of rival generation levels for which a firm's profit-maximising reaction function is undefined. However this does not necessarily imply that an equilibrium does not exist: the regions of generation where reaction functions are defined may indeed contain the intersection.

Given the complicated expressions representing the spot market first order conditions for each firm, it is difficult to find an analytical expression for the reaction functions, especially given the discontinuities implied by the concavity condition. Hence it is even more difficult to define analytical conditions under which an equilibrium exists.

If, for the moment, we were to assume that concave profits were guaranteed, we can describe the likely behaviour of the reaction functions, and thus evaluate the likelihood of an equilibrium. We can then discuss how these conclusions are impacted by regions of convex profits.

Tirole (1988) states that a Cournot-Nash equilibrium exists for two firms, i and j, if:

---

observed, the firm would happily "throw caution to the wind", and attempt to generate at a level which is significantly different from the short-run maximum.

1. The reaction functions for each firm, $R_{i,j}^{t} = \hat{g}_i^t\left(g_j^t\right)$, satisfy the condition $\left|R_{i,j}^{t}{}'\right| < 1$, for every state *t*, where $R_{i,j}^{t}{}' = \frac{d\hat{g}_i^t}{dg_j^t}$. Traditionally, where the goods being traded are strategic substitutes, this slope is negative (i.e., reaction functions are normally downward sloping, but this is not a requirement for an equilibrium).

A sufficient condition for this equilibrium to be unique is:

2. The reaction functions are always concave in the other player's output, i.e, $\frac{d^2 g_i^t}{\left(dg_j^t\right)^2} < 0$. This means that the reaction curve can never turn back on itself and intersect the other reaction function again.

Usually, we can determine the slope of the reaction functions by finding the derivative of the expression for a firm's optimal output as a function of its rival's, with respect to its rival's output. However, in this case, rearranging (11.15) to form an explicit function for the optimal output would be difficult. Instead, we can express the slope of the reaction function as:

$$R_{i,j}^{t}{}' = -\frac{\dfrac{d^2\Pi_i^t}{dg_i^t dg_j^t}}{\dfrac{d^2\Pi_i^t}{\left(dg_i^t\right)^2}} \tag{11.25}$$

Of course, the denominator in 11.25 is the second-order condition for firm *i*'s profit maximising solution, which will be negative for any profit maximising solution to (11.15) . Calculating the numerator in 11.25, we obtain the full expression for the reaction function slope:

$$R_{i,j}^{t}{}' = -\left(\frac{-b \quad +b^2\lambda k_i\left(G^r - E[G] - \left(K - G^r\right)\left(2 - 2\theta_i^r\right)\right)}{-2b - c_i + b^2\lambda k_i\left(G^r - E[G] - \left(K - G^r\right)\left(2 - 2\theta_i^r\right)\right)}\right) \tag{11.26}$$

Since the effect of each firm's output on the contract price is identical, the derivative of the contract price with respect to each firm's output is also identical. 11.26 has been arranged to show this. The slope of the reaction function tends to the traditional Cournot slope, $-b/(2b+c_i)$, as risk aversion, or cost state variability tends to zero. Hence 11.26 shows that the slope of the reaction functions in the destabilisation model is simply the traditional Cournot reaction function slope, with an identical expression added to the numerator and denominator. Since we know that for any reasonable demand and cost curve, the zero-contract Cournot slope satisfies condition (1) above, we can conclude that in the destabilisation case above, $-1 < R_{i,j}'^{\,\prime} < 1$, since the numerator can never grow larger than the denominator when the same value is added to both.

While we do not require it for an equilibrium, we can now attempt to determine if the reaction functions are downward sloping, i.e., expression 11.26 is negative. Since we know that for concave profits, the denominator will be negative, we require that:

$$b \;>\; b^2\lambda k_i\left(G^r - E[G] - \left(K - G^r\right)\left(2 - 2\theta_i^r\right)\right)$$

Or

$$b\lambda k_i\left(G^r - E[G] - \left(K - G^r\right)\left(2 - 2\theta_i^r\right)\right) \;<\; 1 \tag{11.27}$$

Which is even more restrictive than the second order condition, (11.19), which can be expressed as:

$$b\lambda k_i\left(G^r - E[G] - \left(K - G^r\right)\left(2 - 2\theta_i^r\right)\right) \;<\; 2 + \frac{c_i}{b} \tag{11.28}$$

While the second order-condition may be satisfied, the requirement that reaction functions be downward sloping may not. There is a range of low generation levels where both the numerator and denominator of 11.26 are negative, and hence the reaction function slope is downward sloping. There also exists an intermediate range of generation levels where the concavity condition is still satisfied (i.e., the denominator is negative), but the numerator is positive, leading to a positive slope. Over this range, we can be confident that the slopes of the reaction functions are everywhere, in absolute value, less than 1. (As discussed above, it does not make sense to discuss the reaction functions in the case where the denominator of 11.26 is positive).

By condition (2) above, this intersection will be unique if:

$$\frac{d^2 g_i^t}{\left(dg_j^t\right)^2} = \frac{dR_{i,j}^{t\,\prime}}{dg_j^t} < 0 \tag{11.29}$$

Since $R_{i,j}^{t\,\prime}$ is a fractional expression, we can use the quotient rule to evaluate the RHS of 11.26. Let H be the numerator, and J be the denominator, of 11.26, so that:

$$R_{i,j}^{t\,\prime} = -\left(\frac{H}{J}\right)$$

It is clear that:

$$\frac{dH}{dg_j^t} = \frac{dJ}{dg_j^t}$$

Hence we can say that

$$\frac{dR_{i,j}^{t\,\prime}}{dg_j^t} = -\left(\frac{\frac{dH}{dg_j^t}(J-H)}{J^2}\right)$$

$$= -\left(\frac{\frac{dH}{dg_j^t}(b+c_i)}{J^2}\right)$$

$$= \frac{-3b^2\lambda k_i\left(\theta_i^t - 1\right)\left(b + c_i\right)}{\left(-2b - c_i + b^2\lambda k_i\left(G^r - E[G] - \left(K - G^r\right)\left(2 - 2\theta_i^r\right)\right)\right)^2}$$

Since $\theta_i^t \leq 1$, the reaction functions are strictly convex in each other's output, except in the case where firm *i* only faces one cost state (in which case the reaction function will be linear). Hence we cannot say that if an intersection exists, it will be unique. Given that we have shown that the reaction functions are strictly convex, we know that there will be at most two equilibria.

The analysis of the behaviour of the reaction functions, under the assumption that profits are concave and profit maximising solutions to the first order condition exist, provides us with some useful results for analysing their behaviour when this assumption is relaxed.

Firstly, if there exists a solution to the monopoly problem for firm *i,* in a given state, the reaction function for firm *i* is defined at the point at which firm *j* produces no output. From there, as form *j* increases output, the profit maximising reaction function traces a convex curve that has a slope defined by 11.26. The discussion above showed that this slope may be negative or positive, but always, in absolute value, less than 1. Hence, for every unit of extra output produced by firm *j*, total industry output increases by $1+R_{i,j}^{t\prime} > 0$. This moves the concavity bound on total generation closer towards being binding. At some level of firm *j* output, the concavity condition will be violated, and the reaction function will cease to exist at that point, and for any higher generation. If this point occurs before the intersection with firm *j*'s reaction function, no equilibrium will exist. Otherwise, we will observe an equilibrium in that state.

By the same reasoning, if there does not exist a solution to the monopoly problem for firm *i*, there will be no reaction function for that firm, and hence no equilibrium.